IDENTIFYING AND MEASURING COGNITIVE ASPECTS OF A MATHEMATICS ACHIEVEMENT TEST

A Thesis
Presented to
The Academic Faculty

By

Megan E. Lutz

In Partial Fulfillment
Of the Requirements for the Degree
Master of Science in Psychology

Georgia Institute of Technology

May, 2012

Approval Page

Approved by:

Susan E. Embretson
School of Psychology
Georgia Institute of Technology

James Roberts
School of Psychology
Georgia Institute of Technology

Lawrence James
School of Psychology
Georgia Institute of Technology

Date Approved: 26 January 2012

ACKNOWLEDGEMENTS

The author gratefully acknowledges her parents, Michael and Mari Lutz, for their continued support as she wends her way slowly through this process. The author is also grateful to her advisor, Susan Embretson, and her committee for their thoughtful feedback as she worked on this thesis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

Cognitive Diagnostic Models (CDMs) are a useful way to identify potential areas of intervention for students who may not have mastered various skills and abilities at the same time as their peers. Traditionally, CDMs have been used on narrowly defined classroom tests, such as those for determining whether students are able to use different algebraic principles correctly. In the current study, the Deterministic Input, Noisy “And” Gate model (DINA; Haertel, 1989; Junker & Sijtsma, 2001) and the Compensatory Reparameterized Unified Model (CRUM; Hartz, 2002), as parameterized by the log-linear cognitive diagnosis model (LCDM; Henson, Templin, & Willse, 2009), were used to analyze the utility of pre-defined cognitive components in estimating students’ abilities in a broadly defined, standardized mathematics achievement test. The attribute mastery profile distributions were compared; the majority of students was classified into the extremes of no mastery or complete mastery for both the CRUM and DINA models, though greater variability among attribute mastery classifications was obtained by the CRUM.

CHAPTER 1
INTRODUCTION

Assessments of mathematical competency have important consequences for students, teachers, and school systems. Throughout the United States, item response theory (IRT) has become the standard method for calibrating the psychometric properties of items, as well as for estimating a given student's competency level. Various components have been shown to be related to item difficulty in a variety of tests of both mathematical and verbal ability (e.g., Embretson & Daniel, 2008; Gorin & Embretson, 2006; Hornke, 2002). IRT has proven to be a dramatic improvement over classical test theory statistics and analysis, in terms of the information that can be estimated, the usefulness of that information, and the stronger assumptions about the population and the items in question (Embretson & Reise, 2000).

In the realm of mathematics testing, Embretson and Daniel (2008) demonstrated the utility of the Rasch-based logistic linear test models (LLTM; Fisher, 1973) in identifying significant cognitive components for quantitative reasoning items, as well as the advantages gained for using LLTM over simple bivariate regression for identifying such components. In their work, Embretson and Daniel (2008) identified five major cognitive components, which evolved from Mayer, Larkin and Kadane's (1984) analysis: Translation, Integration, Solution Planning, Solution Execution, and Decision Processing. Within each component were attributes, both construct-relevant and construct-irrelevant, that further defined the cognitive components and were significant predictors of Rasch item difficulty. In the case of a mathematics item, a construct-irrelevant attribute would be Contextual Encoding, which is a count of the number of non-mathematical terms in an item. Construct-relevant attributes, then, are those properties of an item that contribute to the item's construct validity, such as Mathematical Encoding, which counts the number of mathematical terms and operators. Mathematics achievement items are written with specific construct-relevant standards and indicators outlined in their blueprint. In the blueprint might be questions about specific geometric shapes and their properties, probabilities of independent events, or properties of different classes of numbers (e.g., Kansas State Department of Education, 2004). In asking about specific geometric shapes, for example, it is necessary to use construct-irrelevant components, such as textual encoding, to even posit the question being asked. Another example of non-standards based cognitive intrusion into an item can occur if a student is given a diagram of a rectangle, containing the dimensions of the rectangle. If the remaining text of the stem asks the student for the perimeter of the rectangle, which the student knows how to do, but the student is unable to translate the diagram to get the dimensions, then the item is made more difficult by the inclusion of the diagram, or the exclusion of the dimensions of the rectangle in text. However tightly the blueprint controls the construct-relevant content of an examination, cognitive components will, by necessity, enter into play, and it is important to know how they can impact both the students' and the items' performance.

Cognitive diagnostic models (CDMs) are an extension of IRT. In the case of dichotomous items, CDMs effectively locate students above or below a mastery threshold for distinct item components. Thus, instead of receiving a raw score on a test, examinees receive a report that further outlines their specific strengths and weaknesses. Traditionally, CDMs have been used on narrowly-defined constructs, where the item attributes are all construct-relevant. An example of such an analysis would be where the attributes are aligned with the standards outlined on a test blueprint, which makes explicit the different aspects of a construct the test is meant to measure. See Huff and Goodman (2007) for further discussion of the utility of cognitive diagnostic assessment.

Within the realm of cognitive diagnosis, there exist several classes of models. For a given item composed of *a* attributes, compensatory models such as the Compensatory Reparameterized Unified Model (CRUM; Hartz, 2002 as cited in Rupp, Templin, & Henson, 2010), an examinee need only master a subset of the *a* attributes in order to correctly solve the item: that is, the mastery of one attribute compensates for the lack of others. The more attributes in the item an examinee has mastered, the better his chances are for solving the item, but the absence of an attribute does not preclude correctly answering an item. With the CRUM, however, no additional advantage is gained for having mastered multiple item attributes: in the context of set theory, it is the union of the item attributes that contribute positively to successful item completion.

In contrast with compensatory models, there are also non-compensatory models, such as the Deterministic Input, Noisy “And” Gate model (DINA; Haertel, 1989; Junker & Sijtsma, 2001), which require mastery of all item attributes for successful item completion: even the absence of only a single attribute should cause the examinee to incorrectly answer the item. In this case, the examinee must have mastery over the intersection of all item attributes.

It is hypothesized that the cognitive components involved in the current study should contribute to an item’s difficulty in a non-compensatory manner, because the construct-irrelevant components of an item all play together towards successfully answering that item – if a student hasn’t mastered one of the cognitive components involved in an item, that should strongly reduce the probability he correctly answers it. That is, for a student to correctly answer a given word problem, he must have mastered not only all of the non-compensatory construct-relevant attributes involved, but also the Translation cognitive component to be able to understand what the item is asking of him in the first place. For the sake of thoroughness, the DINA, CRUM, and the highly flexible LCDM were of interest and used in the diagnosis of construct-relevant and –irrelevant attributes in a state-wide mathematics achievement exam.

CHAPTER 2
METHOD

Participants

The students' data for the present study is a simple random sample of 2,993 examinees from all eighth-graders enrolled in mathematics in the public school system of a Midwestern state. Approximately half (51%) of the students were Male; 74% of the sample were White, 11% were Hispanic, 8% were Black, and 1.5% were Native American. Examinees' responses were scored electronically by the central administrator. Potential identifiers were stripped from the data, and both raw and scored responses were available for analysis.

Instrument

The mathematics exam was developed to be administered in three parts over the course of three days, with each part roughly consisting of the same number of items. A total of 86 multiple-choice items were included, and their classical test theory statistics are in Table 1. All three parts covered the same basic algebraic principles, but there was only slight conceptual overlap among the three sections in terms of the specific principles being tested. For example, Part 1 consisted of 30 items and was the only section in which Pythagorean relationships were tested; Part 2 consisted of 27 items and was the only part that addressed dilation and scaling of objects; Part 3 consisted of 29 items and was where students were tested on their fluency with equations containing two variables.

Table 1. Classical test theory statistics for 86 items.

Item	Attempted	Correct	Proportion Correct	Item-Test Correlation Pearson	Biserial
1	2993	2867	95.8	0.209	0.466
2	2993	2615	87.4	0.301	0.483
3	2993	1704	56.9	0.282	0.356
4	2993	1861	62.2	-0.119	-0.151
5	2993	2423	81.0	0.483	0.697
6	2993	1869	62.4	0.499	0.637
7	2993	2825	94.4	0.212	0.430
8	2993	2628	87.8	0.385	0.622
9	2993	1689	56.4	0.515	0.649
10	2993	1715	57.3	0.353	0.446
11	2993	2507	83.8	0.329	0.494
12	2993	1163	38.9	0.291	0.370
13	2993	1686	56.3	0.365	0.460
14	2993	2780	92.9	0.258	0.488
15	2993	2429	81.2	0.419	0.607
16	2993	2019	67.5	0.386	0.502
17	2993	2406	80.4	0.398	0.572
18	2993	2525	84.4	0.445	0.674
19	2993	1910	63.8	0.494	0.633
20	2993	2155	72.0	0.402	0.536
21	2993	2129	71.1	0.560	0.743
22	2993	2070	69.2	0.541	0.710
23	2993	2347	78.4	0.473	0.664
24	2993	2431	81.2	0.387	0.561
25	2993	1999	66.8	0.479	0.622
26	2993	1304	43.6	0.443	0.558
27	2993	2257	75.4	0.367	0.502
28	2993	1729	57.8	0.430	0.543
29	2993	2249	75.1	0.496	0.676
30	2993	1931	64.5	0.497	0.639
31	2993	2629	87.8	0.300	0.485
32	2993	2293	76.6	0.409	0.565
33	2993	1740	58.1	0.357	0.451
34	2993	2786	93.1	0.318	0.605
35	2993	2261	75.5	0.493	0.674
36	2993	2287	76.4	0.264	0.364
37	2993	2773	92.6	0.368	0.689
38	2993	1618	54.1	0.241	0.302
39	2993	2553	85.3	0.442	0.680
40	2993	2497	83.4	0.453	0.677
41	2993	1952	65.2	0.441	0.568
42	2993	2578	86.1	0.431	0.673
43	2993	2530	84.5	0.350	0.531

Table 1 (continued)

44	2993	1904	63.6	0.532	0.681
45	2993	1886	63.0	0.531	0.680
46	2993	2367	79.1	0.448	0.634
47	2993	1711	57.2	0.413	0.521
48	2993	1493	49.9	0.424	0.531
49	2993	2507	83.8	0.416	0.625
50	2993	1006	33.6	0.502	0.651
51	2993	2828	94.5	0.361	0.739
52	2993	1787	59.7	0.571	0.723
53	2993	1560	52.1	0.509	0.638
54	2993	987	33.0	0.298	0.386
55	2993	1354	45.2	0.412	0.518
56	2993	1633	54.6	0.403	0.506
57	2993	1235	41.3	0.392	0.495
58	2993	2796	93.4	0.290	0.562
59	2993	2692	89.9	0.261	0.446
60	2993	2651	88.6	0.336	0.553
61	2993	2171	72.5	0.375	0.502
62	2993	2501	83.6	0.277	0.415
63	2993	1034	34.5	0.408	0.526
64	2993	2498	83.5	0.451	0.675
65	2993	2769	92.5	0.235	0.438
66	2993	2258	75.4	0.349	0.477
67	2993	1688	56.4	0.402	0.506
68	2993	2009	67.1	0.399	0.518
69	2993	2078	69.4	0.474	0.623
70	2993	1940	64.8	0.446	0.574
71	2993	2033	67.9	0.458	0.598
72	2993	1943	64.9	0.433	0.557
73	2993	2306	77.0	0.280	0.388
74	2993	2056	68.7	0.289	0.378
75	2993	2742	91.6	0.355	0.639
76	2993	2315	77.3	0.374	0.520
77	2993	2635	88.0	0.369	0.599
78	2993	1835	61.3	0.446	0.567
79	2993	1912	63.9	0.469	0.601
80	2993	2535	84.7	0.381	0.580
81	2993	2116	70.7	0.430	0.569
82	2993	2130	71.2	0.417	0.553
83	2993	2445	81.7	0.509	0.743
84	2993	1490	49.8	0.381	0.478
85	2993	2200	73.5	0.363	0.490
86	2993	862	28.8	0.173	0.229

The test forms were developed within the IRT framework, with items written for content validity based on a state-wide blueprint. The blueprint outlined four different achievement domains for eighth-grade students: Number and Computation, Algebra, Geometry, and Data, with several benchmarks in each.

Cognitive Components

Building off of previous research into different stages of processing item components (e.g., Embretson & Daniel, 2008; Gorin & Embretson, 2006; Mayer, Larkin, & Kadane, 1984), five cognitive processes, or components, were identified and defined: Translation, Integration, Solution Planning, Solution Execution, and Decision Processing. Each cognitive component is defined by its cognitive attributes. The attributes within a component are those qualities of an item that comprise a given phase of the problem-solving process (Mayer, et al., 1984). For example, the attributes that define the "Translation" component, including "Mathematical Encoding", "Contextual Encoding", and "Reading Complexity", are all aspects of an item that pertain to taking in and processing the information that is presented. The "Solution Execution" component, as another example, is comprised of attributes that are involved directly in the solving a problem, including "Number of Procedures" and "Number of Computations" necessary for evaluating a problem. A complete list of the attributes and their definitions can be found in Table A1.

The sample item in Figure 1 would be scored in the following way: Encoding = 51 total words and math terms, Mathematical = 33 implicit and explicit math terms, Context = 18 words, Reading Level = 11.0, Recall Knowledge Principles = 1, Bottom-up=1; any remaining attributes not involved in the item are equal to zero. In this way, the component scores in Tables A2 through A5 were developed.

The circumference of a circle is given by the equation $C = 2\pi r$. What is another way to write this relationship? A) * $C = 2(\pi r)$ B) $C = \pi (r \times r)$ C) $C = (\pi \times \pi) r$ D) $C = 2(\pi + r)$

Figure 1. A sample, hypothetical question for eighth-grade items. The correct answer is marked with an asterisk

The identification and subsequent scoring of the cognitive components for the items resulted in a **Q-**matrix of cognitive attributes, with each row in **Q** representing an item. The individual components for each item were assessed by three quantitative psychology graduate students and approved by their advisor, with an average Fleiss' kappa of 0.690 (*min* = 0.394, *max* = 0.953). Cognitive attributes were selected for inclusion in the **Q-**matrix based on their performance in a multiple linear regression of 3PL item difficulties on the cognitive components as well as their zero-order correlations, in which case a minimum correlation of 0.10 was required. In the end, ten attributes were selected for inclusion. The **Q-**matrix used in the analyses can be found in Table 2.

Table 2. Q-matrix columns for full test length and parameter space, with attribute labels

Item Number	Test Position	Number and Computation	Encoding	Contextual Encoding	Encode Diagram	Generate Equations or Plausible Values
1	0	1	0	0	0	0
2	0	1	0	0	0	0
3	0	1	1	1	0	1
4	0	1	1	0	1	0
5	0	1	1	1	0	0
6	0	1	1	1	0	1
7	0	1	0	0	0	0
8	0	1	0	0	0	0
9	0	1	0	1	0	1
10	1	1	0	1	0	0
11	1	1	0	0	0	0
12	1	1	0	1	0	1
13	1	1	0	1	0	0
14	1	1	0	0	0	0
15	1	1	0	0	0	0
16	1	1	1	1	0	0
17	1	1	1	1	0	0
18	1	1	1	1	0	0
19	1	0	1	0	1	1
20	1	0	0	0	1	0
21	1	0	0	0	0	0
22	1	0	0	0	0	0
23	1	0	1	1	1	0
24	1	0	1	1	0	1
25	1	0	1	1	0	1
26	1	0	1	1	1	0
27	1	0	1	1	0	1
28	1	0	1	1	0	1
29	1	0	1	1	0	0
30	1	0	1	1	0	1
31	0	0	1	1	0	0
32	0	0	0	0	0	0
33	0	0	1	1	0	1
34	0	0	0	0	0	0
35	0	0	1	1	0	0
36	0	0	1	1	0	0
37	0	0	1	1	1	0
38	0	0	1	1	1	0
39	0	0	0	1	0	0
40	1	0	1	1	0	0
41	1	0	0	1	1	0
42	1	0	1	1	0	0
43	1	0	0	1	0	0
44	1	0	1	1	1	0
45	1	0	1	1	1	0
	1	0	1	1	1	0

Table 2 (continued)

47	1	0	1	1	1	0
48	1	1	0	1	0	0
49	1	1	0	1	0	0
50	1	1	1	1	1	0
51	1	1	1	1	0	0
52	1	1	0	1	0	0
53	1	1	0	1	0	0
54	1	0	1	1	0	0
55	1	0	1	1	0	0
56	1	0	0	0	0	0
57	1	0	0	1	1	0
58	0	1	1	1	0	1
59	0	1	1	1	0	1
60	0	1	1	1	0	1
61	0	1	1	1	0	1
62	0	1	1	1	0	1
63	0	1	1	1	1	1
64	0	1	1	1	0	1
65	0	1	1	0	0	0
66	0	1	0	0	0	0
67	1	1	1	1	0	1
68	1	1	0	0	0	0
69	1	1	0	0	0	0
70	1	1	0	0	0	0
71	1	1	0	0	0	0
72	1	1	0	1	0	0
73	1	1	0	0	0	0
74	1	1	0	1	0	0
75	1	0	1	0	1	0
76	1	0	0	1	0	0
77	1	0	0	0	1	0
78	1	0	1	1	0	0
79	1	0	1	0	1	0
80	1	0	0	0	1	0
81	1	0	0	0	1	0
82	1	0	0	0	1	0
83	1	0	0	0	0	1
84	1	0	1	1	0	1
85	1	0	1	1	0	0
86	1	0	1	1	0	1

Table 2 (continued)

Item Number	Visualization	Number of Subgoals	Relative Definition of Variables	Bottom-up Processing
1	0	0	0	0
2	0	0	0	0
3	0	0	0	0
4	0	0	0	0
5	1	0	0	0
6	0	0	0	0
7	0	0	0	0
8	0	0	0	0
9	0	0	0	1
10	0	0	0	0
11	0	0	0	0
12	0	0	0	1
13	0	1	0	0
14	0	0	0	0
15	0	0	0	0
16	0	0	0	0
17	0	0	0	0
18	0	0	0	0
19	0	0	0	1
20	0	0	0	0
21	1	0	0	0
22	0	0	0	1
23	0	0	0	0
24	0	0	0	1
25	0	0	0	0
26	1	0	0	1
27	0	1	0	0
28	0	1	0	0
29	0	0	0	1
30	0	1	0	0
31	0	0	0	0
32	0	0	0	0
33	1	0	1	0
34	0	0	0	0
35	0	0	0	0
36	0	0	0	1
37	0	0	0	0
38	0	0	0	0
39	0	0	0	0
40	0	0	0	0
41	0	1	0	0
42	0	0	0	0
43	1	0	0	0
44	0	0	0	0
45	0	0	0	0

Table 2 (continued)

46	0	0	0	0
47	0	0	0	0
48	0	1	0	0
49	0	0	0	0
50	0	0	0	0
51	0	0	0	0
52	1	0	0	0
53	1	0	0	0
54	0	0	0	0
55	0	0	0	0
56	1	0	0	0
57	0	1	0	0
58	0	0	0	0
59	0	0	0	0
60	0	0	0	0
61	0	0	0	0
62	0	0	0	0
63	0	0	0	0
64	0	0	0	0
65	0	0	0	0
66	0	0	0	0
67	0	0	0	1
68	0	0	0	0
69	0	0	0	0
70	0	0	0	0
71	0	0	0	1
72	0	0	0	0
73	0	0	0	1
74	0	0	0	1
75	0	0	0	0
76	0	0	0	1
77	0	0	0	1
78	1	0	1	1
79	0	0	0	0
80	0	0	0	0
81	0	0	0	0
82	0	0	0	1
83	0	0	0	0
84	1	1	0	0
85	0	1	0	1
86	0	0	0	0

It is important to note that, as the three models used for the analysis are classification models, they identify examinees as mastering or not mastering an attribute based on some threshold probability of mastery ($p = 0.50$), and not located along a mastery scale as in IRT; the **Q-**matrix reflects this dichotomy by recoding the non-binary cognitive components into binary indicators, based on a median split for inherently continuous components like “Contextual Encoding” and “Number of Computations”, or based on presence or absence of a component, such as “Relative Definition of Variables”. Both the CRUM and DINA were fit using the same cognitive components and **Q**-matrix.

Models of Interest

LCDM

The LCDM is a general CDM, in that it can estimate a wide range of common and uncommon CDMs. The most general form of the LCDM is outlined in Equation 1:

$$\Pr\left[X_{ij} = 1 \mid \boldsymbol{\alpha}_j, \mathbf{q}_i\right] = \frac{\exp\left\{\boldsymbol{\lambda}_i^T \mathbf{h}\left(\boldsymbol{\alpha}_j, \mathbf{q}_i\right) - \eta_i\right\}}{1 + \exp\left\{\boldsymbol{\lambda}_i^T \mathbf{h}\left(\boldsymbol{\alpha}_j, \mathbf{q}_i\right) - \eta_i\right\}} \tag{1}$$

where $\boldsymbol{\alpha}_j$ is a vector containing the attribute mastery pattern for student *j*, $\mathbf{q}_i$ is the attribute pattern for item *i*, $\boldsymbol{\lambda}_i$ is a vector of weights for item *i*, and $\mathbf{h}(\cdot)$ is a set of linear combinations of examinee and item attribute patterns; depending on $\mathbf{h}(\cdot)$, an item can be conjunctive (multiplicative) or compensatory (additive) in nature. Finally, η_i is the base-line probability of getting the item correct for the reference group, which is that group of students that has not mastered any attributes (i.e., $\boldsymbol{\alpha}_j = \mathbf{0}$ for all students in the reference group). The LCDM is sufficiently general such that the DINA and CRUM, as well as other common cognitive diagnostic models, are nested within it, which facilitates model and parameter comparisons, as well as model estimation.

Although the LCDM is a general—and, theoretically, flexible—CDM, the current state of the art does not allow practitioners to specify custom item models. That is, besides pre-existing models (e.g., DINA, CRUM), one cannot at this time specify the level or nature of interactions estimated in the LCDM without writing one's own software. A preliminary run of the LCDM model indicated limited utility of the final results; indeed, the full model would contain a total of 1,270 item parameters, most of which may not be theoretically interesting, and many of which failed to converge. The complexity of the item-side of the model also makes estimation and interpretation of the person-side attribute mastery patterns a daunting task. While it is desirable to investigate a model of the LCDM form, specifically the model containing only the highest-level interaction terms and main effects, the current state of the estimation software does not allow for such customizability, and so the LCDM was not included in the final analysis for the current study.

DINA

As cognitive components for math items are thought to be non-compensatory, the use of the DINA model is appropriate. In the context of LCDM, the DINA arises when $\boldsymbol{\lambda}_i^T$ is fixed to zero for all $\mathbf{h}(\boldsymbol{\alpha}_j, \mathbf{q}_i)$ not involving the highest order interaction. For an item containing both of two attributes, e.g., $\mathbf{q}_i = \begin{bmatrix}1 & 1\end{bmatrix}$, the only non-zero elements in $\boldsymbol{\lambda}_i^T$ are those for the intercept of the item and the interaction of attributes 1 and 2. Thus, Equation 2 gives the probability of examinee *j* correctly answering such an item:

$$\Pr[X_{jc} = 1 | \alpha_{c1}, \alpha_{c2}] = \frac{\exp\{\lambda_{i,0} + \lambda_{i,2,(1,2)}\alpha_{c1}\alpha_{c2}\}}{1 + \exp\{\lambda_{i,0} + \lambda_{i,2,(1,2)}\alpha_{c1}\alpha_{c2}\}} \tag{2}$$

Equation 3 shows the original parameterization of the DINA, based on slip (*s*) and guess (*g*) parameters, which are, respectively, the probability of answering an item incorrectly when one has mastered the necessary attributes and the probability of answering an item correctly when one has not mastered all required attributes.

$$\Pr[X_{ij} = 1 | \boldsymbol{\alpha}, s_j, g_j] = (1 - s_j)^{\xi_{ij}} (g_j)^{1-\xi_{ij}};$$

$$\xi_{ij} = \begin{cases} 0 & \text{examinee has not mastered all attributes} \\ 1 & \text{examinee has mastered all attributes} \end{cases} \tag{3}$$

Thus, examinee *j* has probability (1-s_j) of answering item *i* correctly when he has mastered all of the attributes required, and probability g_j of answering item *i* correctly when he has failed to master all of the attributes required (Junker and Sijtsma, 2001). Fairly straightforward mathematics allows one to calculate the DINA parameters from the LCDM representation. Recalling the definitions of the different model parameters, g is the conditional probability of correctly answering an item when an examinee does not possess all required attributes. In the LCDM framework, then, g is related to the intercept (Equation 4). Similarly, s is related to the interaction in the LCDM parameterization (Equation 5).

$$g = \frac{\exp\{\lambda_{i,0}\}}{1+\exp\{\lambda_{i,0}\}} \tag{4}$$

$$s = 1 - \frac{\exp\{\lambda_{i,0} + \lambda_{i,2,(1,2)}\alpha_{c1}\alpha_{c2}\}}{1+\exp\{\lambda_{i,0} + \lambda_{i,2,(1,2)}\alpha_{c1}\alpha_{c2}\}} \tag{5}$$

CRUM

Although the cognitive components in the current study are theoretically non-compensatory, due diligence in model comparisons suggests investigating the other extreme: that of a completely compensatory relationship of item attributes. In the LCDM parameterization, the CRUM model takes the form of Equation 6, which involves only the main effects of each attribute. That is, $\boldsymbol{\lambda}_i^T$ is fixed to 0 for all $h(\mathbf{a}_j, \mathbf{q}_i)$ involved in any interaction among item attributes. For an item containing both of two attributes, e.g., $\mathbf{q}_i$ = [1 1], the only non-zero elements in $\boldsymbol{\lambda}_i^T$ are those for the intercept of the item and the main effects of attributes 1 and 2. Thus, the probability that examinee j correctly answers such an item is provided by Equation 6.

$$\Pr[X_{ij} = 1 | \alpha_{j1}, \alpha_{j2}] = \frac{\exp\{\lambda_{i,0} + \lambda_{i,1,(1)}\alpha_{j1} + \lambda_{i,1,(2)}\alpha_{j2}\}}{1+\exp\{\lambda_{i,0} + \lambda_{i,1,(1)}\alpha_{j1} + \lambda_{i,1,(2)}\alpha_{j2}\}} \tag{6}$$

The LCDM parameterization of the CRUM looks very similar to the original drafting of the CRUM model (Equation 7).

$$\Pr[X_{ij} = 1 \mid \alpha_{j1}, \alpha_{j1}] = \frac{\exp\left\{\lambda_{i,0} + \lambda_{i,1,(1)}\alpha_{j1}q_{i1} + \lambda_{i,1,(2)}\alpha_{j2}q_{i2}\right\}}{1 + \exp\left\{\lambda_{i,0} + \lambda_{i,1,(1)}\alpha_{j1}q_{i1} + \lambda_{i,1,(2)}\alpha_{j2}q_{i2}\right\}} \tag{7}$$

where q_{ia} represents whether attribute a is present on item i (Rupp, Templin, and Henson, 2010).

Model Estimation and Specification

The estimation of the models was performed using Markov Chain Monte Carlo (MCMC) methods (Henson, Templin, & Willse, 2009), specifically Metropolis-Hastings within Gibbs sampling. The estimation was carried out via a custom FORTRAN program (as described in Henson, Templin, and Willse, 2009). Each model was estimated using a single chain consisting of 20,000 steps, of which the first 17,000 were used to "burn in" the parameter estimates, resulting in a chain of 3,000 estimates for each parameter in the model. The item parameters $\boldsymbol{\lambda}_i^T$ and η_i started with $U(-10,10)$ prior distributions for all $i = 1, 2, \ldots, I$. The priors for the person parameters, $\boldsymbol{\alpha}_j$, however, are empirically derived due to the possibility of inter-attribute correlations, and so are assumed to follow a dichotomized multivariate normal distribution (Henson, Templin, & Willse, 2009), which reflects the binary classification of examinees on each of $K = 10$ attributes.

Model Convergence and Comparison

MCMC methods, being iterative in nature, require the researcher to investigate the parameter chains for convergence of the estimates in the post burn-in phase of the process. One common convergence assessment for a single chain of estimates is the Geweke convergence diagnostic (Geweke, 1992), which is, in essence, a two-sample *t*-

test with a correction in the standard error for autocorrelation of the variables. Because of high auto-correlation in the current study, the Geweke statistic could not be calculated using the usual proportions of 10% and 50% for comparison: the smallest proportions that could be consistently calculated for the models was 40% and 40%, meaning that 2,400 cases were used in calculating the statistic, increasing the power to find small differences. The average absolute change in the parameters for the full DINA and CRUM models was 0.026 and 0.095, respectively. Alternative methods for assessing convergence include graphical techniques, such as trace plots of the parameter chains and regression plots; numerical techniques, including correlation and regression analysis of the Geweke-window parameter means; statistical techniques, including the use of standardized differences of the Geweke-window means to flag problematic parameters. The impact of non-convergent parameters on predicted probabilities can also be investigated for practical significance.

The reduced models used for comparison were reduced in both the parameter space and the test length. The **Q-**matrix was reduced to contain the eight cognitive attributes with the highest representation among the items. No examinees were excluded in the reduced model analyses, and both the DINA and CRUM calibrations on the reduced dataset were run and compared using the same methods as on the full datasets. The results of the reduced models are contained in Appendix C.

As the DINA and CRUM models are not nested, being on either end of the compensatory spectrum, a chi-squared test for improved fit cannot be conducted for the two of them. Instead, relative fit may be compared via other traditional methods, such as AIC, BIC, and comparison of log-likelihoods.

CHAPTER 3: RESULTS

Models Used

LCDM

The preliminary run of the full LCDM took 21.5 days to run to completion; the reduced LCDM ran in just over 4 days. Discussion of the results at this point will be confined to those of the full LCDM model, using all 86 items and 10 cognitive attributes. Including the intercepts, a total of 1,270 item parameters were estimated. The inherent complexity of the full LCDM model in this case illustrates that the current state of LCDM estimation falls short for large **Q** matrices or long tests, with parameter lists quickly getting out of control. In the absence of a more customizable program, in which once can reduce the complexity of each item model to include, say, only main effects and highest-order interactions, estimation of a non-standard version of the LCDM is infeasible.

DINA

The full DINA took 2.5 days to run to completion; the reduced DINA ran in one day. Discussion of the results at this point will be confined to those of the full DINA model; supplemental results for the reduced DINA can be found in Appendix C. Figure 2 illustrates that, for all possible attribute mastery patterns, the majority of students are classified as either possessing all attributes or only one of them.

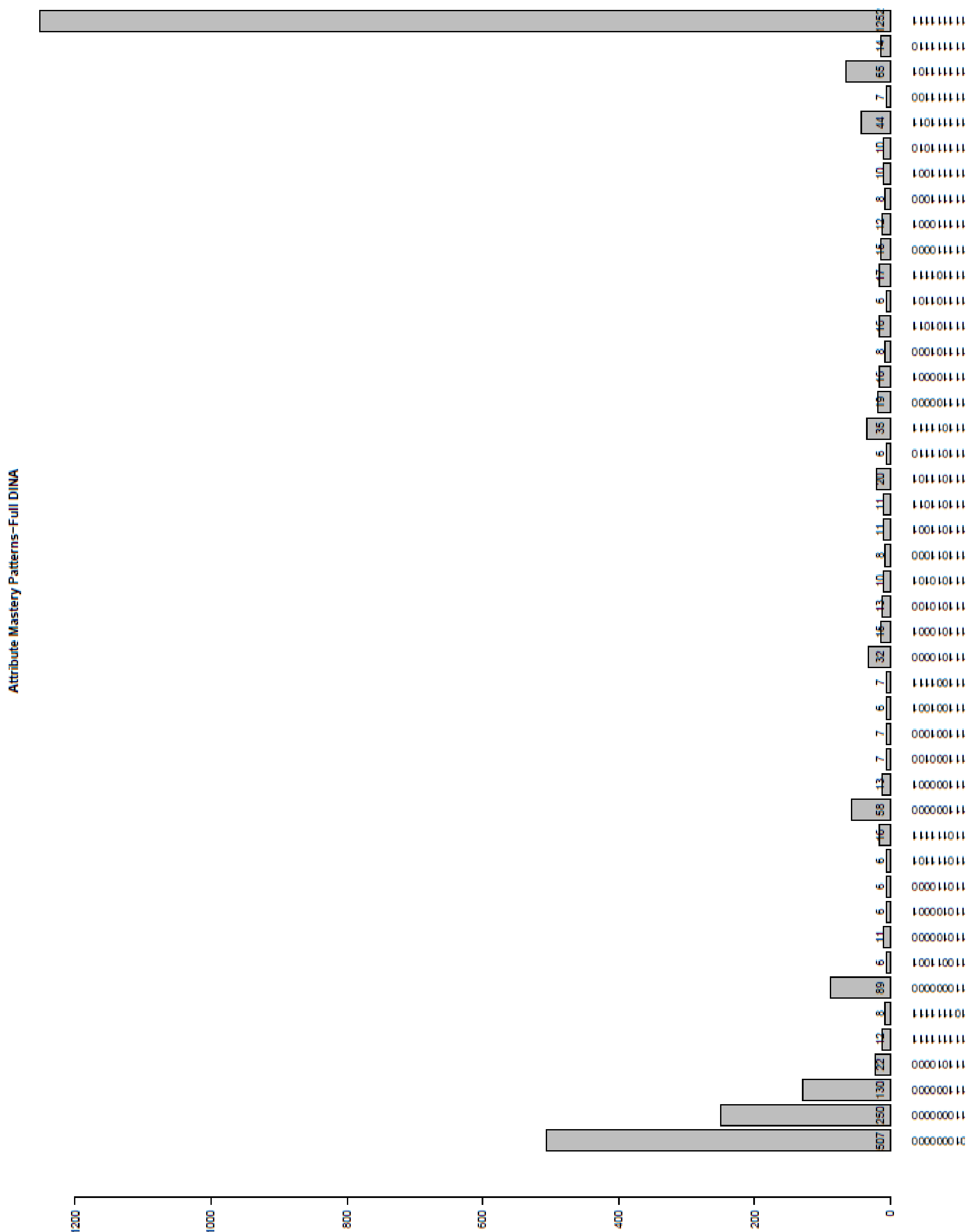

Figure 2. *A plot of the distribution of DINA-estimated attribute mastery profiles. The two extremes of non-mastery and total-mastery are over-represented, while few students have been estimated to have the middle profiles.*

Table 3 shows the four most populous patterns and their frequencies for the DINA, which account for 71.5% of the total sample.

Table 3. Most common DINA classification patterns

Pattern	Frequency (%)
[1111111111]	1,252 (41.8)
[0010000000]	507 (16.9)
[0110000000]	250 (8.4)
[0111000000]	130 (4.3)

The distribution in Figure 2 arose from the attribute probabilities estimated by the FORTRAN program, outlined in Table 4. If an examinee had an attribute mastery probability greater than 0.50, then that student was determined to have mastered the attribute, represented by a ‘1’ in the attribute mastery pattern; with a probability less than 0.50, the student was determined not to have mastered the attribute, represented by a 0 in the attribute mastery patterns of Table 3 and Figure 2.

Table 4. Attribute mastery probabilities for DINA: Examinees 1-15

Student	Test Position	Number and Computation	Encoding	Contextual Encoding	Encode Diagram	Generate Equations or Plausible Values
1	1	1	1	1	1	1
2	1	1	1	1	1	1
3	1	1	1	1	1	1
4	0.597	0.864	0.907	0.154	0.07	0.399
5	1	1	1	1	1	1
6	1	1	1	1	1	1
7	1	1	1	1	1	1
8	0.069	0.001	0.666	0.001	0.007	0.021
9	1	1	1	1	1	1
10	0.052	0.134	0.672	0.001	0.035	0.043
11	1	1	1	1	1	0.962
12	0.042	0.285	0.685	0.001	0.054	0.065
13	1	1	1	0.999	0.855	0.723
14	0.982	1	0.991	0.923	0.227	0.945
15	1	1	1	1	0.999	0.995

Table 4 (continued)

Student	Visualization	Number of Subgoals	Relative Definition of Variables	Bottom-up Processing
1	1	1	0.882	1
2	1	1	0.981	1
3	1	1	0.818	1
4	0.023	0.138	0.132	0.076
5	0.972	0.012	0.537	0.999
6	1	1	0.979	1
7	0.999	0.986	0.973	1
8	0	0.003	0.012	0.001
9	1	0.999	0.98	1
10	0.005	0.017	0.031	0.022
11	0.985	0.988	0.776	0.821
12	0.014	0.021	0.066	0.037
13	0.789	0.132	0.259	0.312
14	0.018	0.863	0.331	0.048
15	0.954	0.918	0.371	0.991

CRUM

The full CRUM took five days to run to completion; the reduced CRUM ran in 1.5 days. Discussion of the results at this point will be confined to those of the full CRUM model; supplemental results for the reduced CRUM can be found in Appendix C. Like Figure 2 for the DINA, Figure 3 shows that the four most common attribute mastery patterns from the CRUM still have most of the cases diagnosed into the extremes of complete mastery or complete non-mastery, but with more variability among the remaining cases.

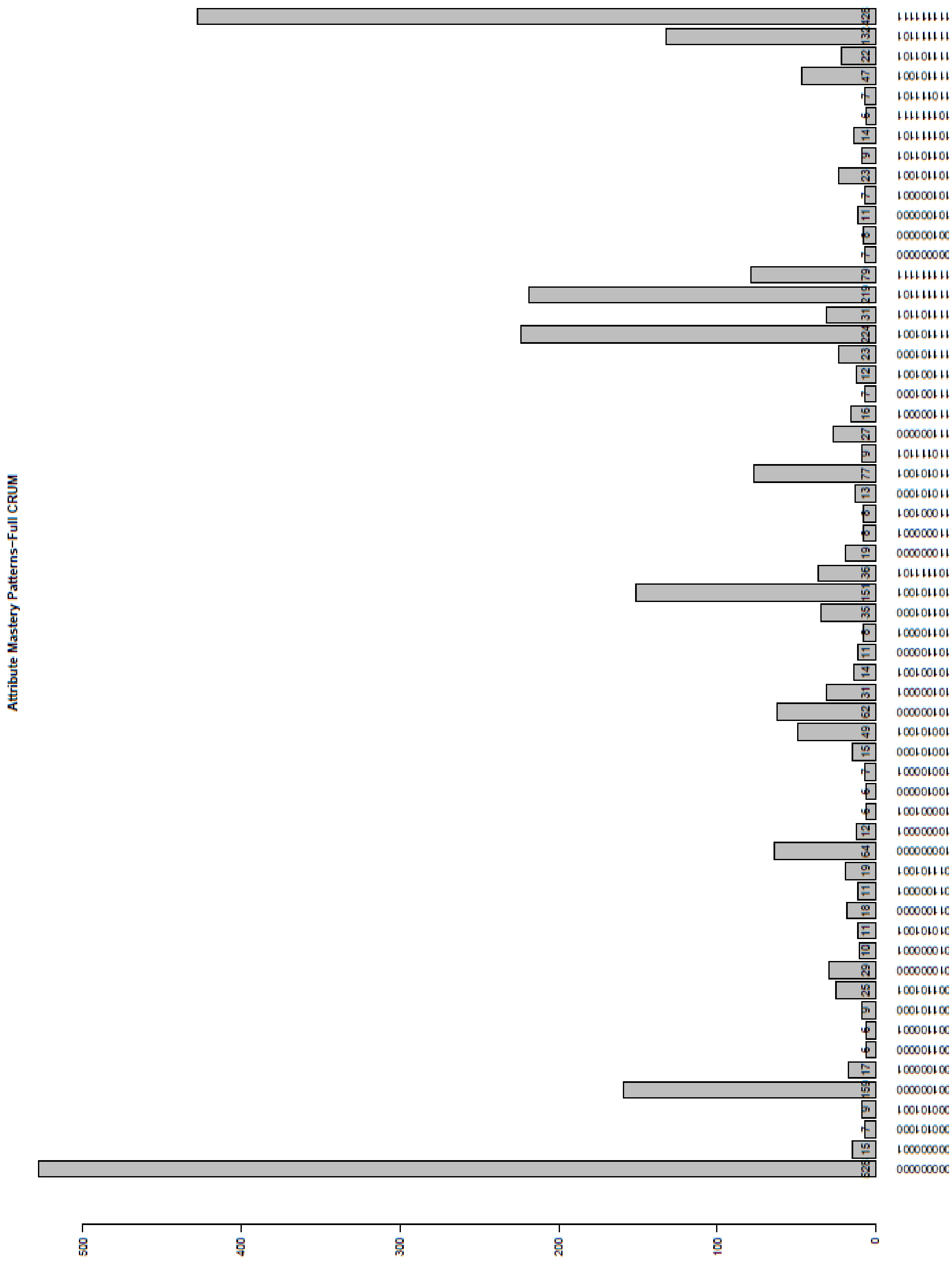
Attribute Mastery Patterns–Full CRUM

Figure 3. A plot of the distribution of CRUMimated attribute mastery profiles. The two extremes of non-mastery and total-mastery are highly-represented, but there is a broad distribution of students among the less extreme profiles as well.

Indeed, the next most common mastery patterns indicate more mastery than not. Table 5 outlines the four most populous patterns and their frequencies for the CRUM, which account for 46.7% of the total sample. Table 6 lists the CRUM-estimated attribute mastery profiles for the first 15 examinees.

Table 5. Most common CRUM classification patterns

Pattern	Frequency (%)
[0000000000]	528 (17.6)
[1111111111]	428 (14.3)
[0111101001]	224 (7.5)
[0111111101]	219 (7.3)

Table 6. Attribute mastery probabilities for CRUM: Examinees 1-15

Student	Test Position	Number and Computation	Encoding	Contextual Encoding	Encode Diagram	Generate Equations or Plausible Values
1	0.994	1	0.952	0.995	1	0.963
2	0.004	0.844	0.729	0.056	0.034	0
3	0	1	0.025	1	1	0.734
4	0.981	1	0.993	0.997	1	0.829
5	0	0.988	0.969	0.998	0.999	0.794
6	0	0.003	0	0	0	0
7	0.337	0.99	0.953	0.993	0.99	0.658
8	0.002	0.004	0.021	0.038	0	0
9	0.068	0.996	0.901	0.833	0.996	0.063
10	0	0.001	0	0	0.015	0
11	0.184	0.964	0.294	0.676	0.969	0.058
12	0	0.935	0.885	0.97	0.044	0
13	0.358	0.993	0.712	0.769	0.97	0.28
14	0	0.001	0.002	0.002	0	0
15	0.571	0.971	0.049	0.939	0.212	0.142

Table 6 (continued)

Student	Visualization	Number of Subgoals	Relative Definition of Variables	Bottom-up Processing
1	1	0.964	0.217	0.995
2	0	0	0	0.037
3	1	0.494	0	0.988
4	1	0.892	0.673	0.985
5	1	0.754	0.444	0.995
6	0	0	0	0.021
7	1	0.692	0.26	0.998
8	0	0	0	0.013
9	1	0.421	0.001	0.902
10	0	0	0	0.004
11	0.983	0.063	0	0.687
12	0	0	0	0.008
13	0.97	0.47	0	0.987
14	0	0	0	0.014
15	0.425	0.158	0	0.265

Model Convergence

Convergence was assessed by various analyses of the first 40% and last 40% of the post-burn-in steps of the MCMC chain. For example, a scatter plot of the averages of the DINA slip parameters (Figure B13) from the first 40% and last 40% of the chain indicates near-perfect linearity; a follow-up simple linear regression analysis of the same averages yields Equation 8.

$$slip.first = .992 * slip.last \tag{8}$$

An investigation of the standard errors of the regression parameters (Table B1) reveals that Equation 8 is not statistically different from the line $y = x$ and, therefore, one can conclude convergence of the slip parameters, on average. Similar regression results were obtained for the guess parameters of the DINA, as well as the main effects for the CRUM, in all cases failing to reject the hypothesized model $y = x$, or that of convergence of the average model parameters.. In addition to the traditional Geweke convergence diagnostic (Geweke, 1992), the current study also utilized graphical and other statistical techniques, which are outlined in this Appendix. All convergence analyses were conducted on the post-burn-in parameter chains of the MCMC analysis, which entailed the final 3,000 steps of the chains.

Graphical Techniques

A common, first-glance assessment of model stability is the inspection of parameter trace plots, which track the movement of the parameter estimate. A stable model will have parameters with variability around a horizontal line, the mean of the parameter chain, which is therefore representative of the final parameter estimate. If a trend is apparent in the trace plot, one can infer a change in the mean of the estimate over time, which means the parameter is not stable and that the model might not have converged. A prototypically "good" trace plot, for an apparently stable parameter, is shown in Figure B1; correspondingly "bad" trace plots are shown in Figures B2-B12. One can see in Figures B2-B12 that the items have a non-random pattern in one of their parameters, whereas the parameter profiled in Figure B1 is pretty stable, with random oscillations about the grand mean.

Numerical techniques

In addition to trace plots, scatter plots, correlation analysis, and regression analysis of the averages from the two Geweke windows in a parameter chain can reveal apparent convergence issues as well as items that may be problematic. By necessity, trace plots take the chain parameter-by-parameter, which in the case of the 371-parameter CRUM can be a burdensome task for the researcher to undertake. Regression and scatter plots, however, provide a holistic glance by looking at which pairs of parameter averages are different, via a parameter's large residual in the plot. Figures B13-B16 illustrate the linear relationship, respectively, of the DINA slip and guess parameters, the CRUM intercepts, and the CRUM main effects, with respective correlations of 0.999, 0.999, 0.999, and 0.945. Correlations, especially correlations of averages, are not enough to conclusively determine whether a model has converged; neither is a scatter plot sufficient to determine the nature of the relationship between the two chain windows, in that one needs the actual equation to know how the estimates line up. Regression analysis of the two windows put a form to the scatter plot, and Tables B1-B4 contain the ANOVA results for the regression analyses for the four sets of parameters—the DINA slip and guess parameters, the CRUM intercepts, and the CRUM main effects—as well as the coefficient estimates and standard errors. Regression analysis indicates convergence when the estimated equation is not statistically different from the line $y = x$, which was the case for each of the four parameter types. However, due to the use of averages for the data points, this technique is relatively insensitive to lack of convergence for particular parameters.

Statistical techniques

The final method of assessing convergence is of the same form of the Geweke statistic, but uses the pooled standard deviation of the same parameter chain windows instead of the respective spectral densities to estimate the standard deviation of the difference, which yields the standardized difference, as in Equation 9.

$$stdiff = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_{pooled}}} \tag{9}$$

The standardized difference can be interpreted the same way as Cohen's D, which is to say it is a directional effect size. Figures B17 – B20 show the distribution of the DINA slip and guess parameters and the CRUM intercept and main effects parameters. As one may expect from the regression analysis and scatterplots, a few parameters fall outside of the $N(0,1)$ 67% confidence limits—"flagged" items—with the largest standardized differences occurring in the CRUM: none of the standardized differences is greater than 3. Tables B5-B7 contain the standardized differences for the DINA parameters and for the CRUM intercept and main effects parameters, respectively. One can see that, for the DINA model, only one parameter of the 172 parameters estimated has a standardized difference outside the 67% normal confidence limits; similarly, only 10 of the 371 CRUM parameters fall outside of the same confidence limits, for respective rates of nonconvergence of 0.5% and 2.7%. Figures B2-B12 provide another look at the 11 questionable parameters flagged by "large" standardized differences, and contain the trace plots for each. The trace plots for the flagged parameters can be compared to the prototypically "good" trace in Figure B1 to help identify periodicity or other non-random patterns in the chain.

Impact of autocorrelation on numerical and statistical methods

There is one limitation to be noted when using the standardized difference for flagging potentially bad items, and that is the inherent autocorrelation of MCMC parameter chains. One way to work around the situation, which was unavailable for the current study, is to thin the post-burn-in chain at a lag determined to have sufficiently low autocorrelation and use estimates and conduct the subsequent analyses based on that smaller chain. As the state of the modeling program does not currently allow the user to thin the parameter chains, the other option is to continue, as in the current study, and calculate the pooled standard deviation in the normal way, being aware that autocorrelation has an inflationary impact on the variance of the estimator, according to Equation 10,

$$s^2\{\bar{\theta}\} = \frac{\sigma^2}{n}\left(\frac{1+\rho_k}{1-\rho_k}\right) = \frac{\sigma^2}{n}SSIF \qquad 10$$

making the detection of small absolute differences more difficult (Walsh, 2004). In Equation 10, ρ_k is the autocorrelation at lag *k*, and *SSIF* is the sample size inflation factor, which indicates how many times bigger a sample size must be to obtain the same amount of precision as an uncorrelated sample. In the case of the DINA model, the maximum autocorrelation at lag 1 was for item 4, at $\rho_1 = 0.998$, meaning the post burn-in parameter estimates were only as precise as those obtained by an uncorrelated sample of size 2. The lone flagged parameter in the DINA estimates was that for Item 8, which had the fourth-largest $\rho_1 = 0.961$ and *SSIF* = 50.28. An argument can be made that the three higher autocorrelations could mask instability at least as bad as that observed for Item 8. Inspection of the three trace plots indicates that, visually, this might have been the case for only one item, item 66. The standardized difference for the guess parameter of item 66 is 0.8002, which can be considered close to the 67% confidence limit of 0.97 when one considers the inflation of its denominator.

Perhaps not unsurprisingly, the CRUM model fared worse, likely due to the dramatically larger number of parameters involved; 20% of the post-burn-in parameter chains had a 1st-order autocorrelation exceeding 0.99, or an *SSIF* no smaller than 199. Thus, it is possible and likely that some standard differences were artificially diminished, making detection of truly unstable parameter estimates more difficult in those cases. The effect of an autocorrelated parameter chain on attribute mastery classification in the current study is unknown.

Practical techniques

It is not enough to simply flag parameters as either stable or non-stable: what may be statistically significant may not be of practical significance. To aid in the interpretation of what the difference in a flagged parameter's averages means for the item, one can look at the difference in the probability of successful item completion, by holding all other parameters constant and changing only the one of interest. In this way, one can tease out practical whole-item information from its individual parameters. The calculations for the difference in probability were conducted using the assumption of complete attribute mastery on the student side for all attributes involved in the item. Table B8 contains the probabilities for successful item completion for the first and last portions of the chain; the probabilities for successful item completion for the overall chain, which is based on the grand mean of the parameters and is in line with the final parameter estimates from the program; and the difference in the probabilities between the first and last portions of the chain. One can see that, with the lone exception of item 34, the apparent instability of all flagged parameters contributed to a change in the probability of correctly answering their associated items by less than 1%.

Convergence discussion

A worthy follow-up to the convergence analysis would investigate the potential causes of instability among the small proportion of flagged items, which may include an extreme item facility, or a very easy item; extreme attribute saturation of the containing item, or an item containing many attributes; and the attribute's representation across items on the whole examination. For the flagged items in the two models, the three potential causes are outlined in Table B9. Based on Figure B21, only item 78 appears unusual in its attribute saturation; the remaining attribute counts are fairly well-represented throughout the 86 items. The attributes' test representation also does not appear to be an indicator of parameter instability: only attribute 9, on item 78 again, was poorly represented throughout the test, with the remainder of the attributes' being present on at least 9 items. In the case of the flagged item intercepts, one cannot look at attribute representation because, by definition, the intercept is a separate estimate from the LCDM-estimated parameters. One may be able to infer that, low attribute saturation (e.g., items with only one or two attributes present) may have less stable intercepts, but as only

6 such items were flagged in the current study, out of a possible 24, that may just be conjecture. Finally, facility does not appear to play a major role in the instability of the parameters. Although items 38, 48, and 56 have relatively low facility, item 51 has fairly high facility, and the remaining 7 items are all in middle range, neither grossly difficult or ridiculously easy. To this end, the DINA and CRUM were estimated again using a reduced item- and attribute-space.

Reduced Parameter Space Analysis

In order to investigate the impact of a large parameter space, both models were estimated on a reduced data set, containing 61 of the original 86 items and 8 of the original 10 attributes. Items were eligible for selection by having a facility of below 0.90 and a sufficiently high biserial correlation of at least 0.386. The attributes were chosen based on their test representation and theoretical viability: Relative Definition of Variables was only present on 2 of the 86 items; Encoding was represented on 33 items, but is partially measured by, and is therefore correlated with, Contextual Encoding, which had a stronger zero-order correlation with item difficulty. Table C1 provides the descriptive statistics for the 61 selected items' facility and biserial correlations. The reduced **Q**-matrix, augmented with the item numbers, is provided in Table C2, along with attribute labels. The same convergence analyses were conducted on the reduced item set as for the full set, with ANOVA results contained in Tables C3-C6 and standardized differences in Tables C7-C9. Figures C1-C4 are a graphical illustration of how tightly the first and last Geweke-window averages of the parameter chains fit the line $y = x$, once again indicating reasonable convergence of the means. Again, the regression analysis is only one piece of evidence for convergence of the parameters on average, and cannot be taken by itself as an indicator of overall model convergence.

One can see that the simultaneous reduction in computational load and improved item and attribute selection yielded marginally better results in terms of convergence, as evidenced by fewer flagged parameters. Tables C10 and C11 further breakdown the flagged parameters, all of which originated in the CRUM estimation, indicating almost uniformly small changes in probability of successfully answering an item (Table C10) and no standouts in terms of test representation or attribute saturation (Table C11 and Figure C12). Tables C12-C14 provide the parameter summaries for the DINA slip, guess, and CRUM parameters. Trace plots of the flagged parameters are provided in Figures C5-C11.

The regression analyses and diagnostics in Tables C3-C6 and Figures C1-C4 all indicate that a linear relationship between the first window and last window parameter means is appropriate and that, furthermore, the relationship is not significantly different from $y = x$. However, the use of the standardized difference for flagging unstable parameters identified different items (Tables C7-C9); only item 78 was flagged in both analyses, for its CRUM main effects. However, item 78 was flagged in the full set for Relative Definition of Variables, which was not included in the reduced analysis: furthermore, item 78 was flagged twice in the reduced set, which may be more of an indication that it is a problematic item in general and not necessarily due to its attribute profile.

No major conclusions can be drawn by the comparison of the full and the reduced parameter and item sets. If one is interested in investigating the impact of an unstable parameter in a large parameter space, or of a poorly performing item, then one is better off reducing the parameter space on a smaller scale than the present study: a large reduction of over a quarter of the items in the current situation might have masked potential improvements in some parameters that would be better revealed by a smaller reduction.

Model Selection

Due to the apparent convergence of the full DINA and full CRUM models, they were retained over the reduced versions of each model. Table 7 shows the relative fit indices for the DINA and CRUM estimates; the CRUM has smaller AIC, BIC, and log-likelihood, indicating it fits the data better than the DINA.

Table 7. Relative model fit indices for DINA and CRUM

Model	Log-Likelihood	Chi-Squared (*k*)	AIC	BIC
DINA	-144,830	289,660.7 (537)	290,734.7	293,958.8
CRUM	-119,952	239,904.2 (371)	240,646.2	242,873.6

Table 8 summarizes the Mean Absolute Deviation (MAD) and Root Mean Squared Error (RMSE) for item difficulty for the two models, which indicates, again, better fit for the CRUM, as there is lower item-level misfit on average for the CRUM than for the DINA.

Table 8. Item-level misfit indices for DINA and CRUM

Model	Mean Absolute Deviation	Root Mean Squared Error
DINA	0.1098	0.1439
CRUM	0.0467	0.0626

Figures 4 and 5 provide a measure of absolute model fit by plotting observed and model-predicted proportions of examinees versus total score on the exam for the DINA and CRUM, respectively All of the relative and absolute fit indices outlined in Tables 7 and 8, as well as the plots in Figures 4 and 5 indicate that CRUM has better model-level and item-level fit than the DINA, which refutes the original hypothesis that a completely non-compensatory model, such as the DINA, would best model the cognitive components.

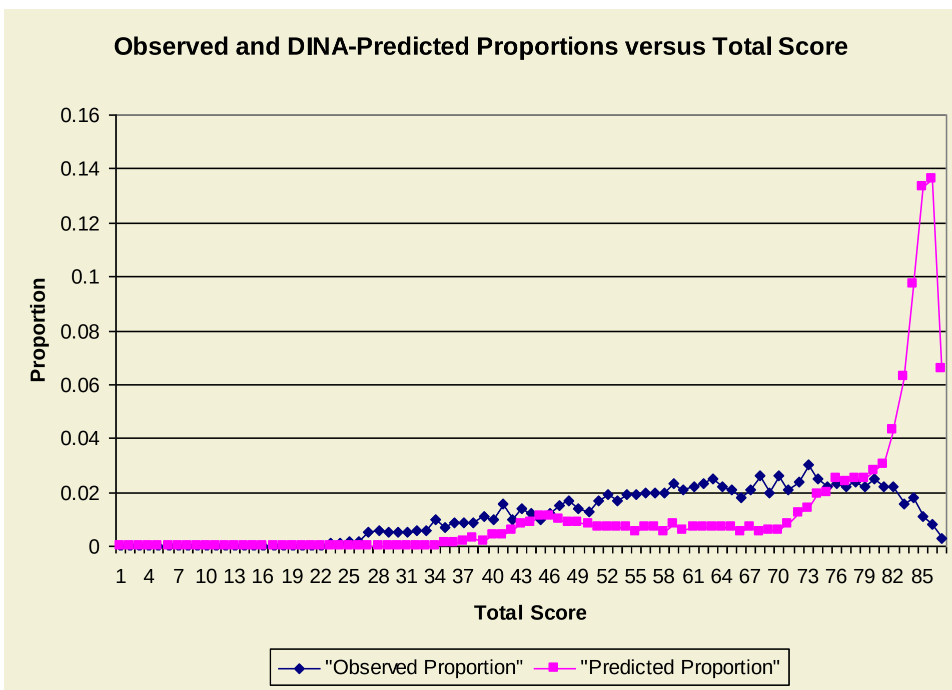

Figure 4. A plot of observed and predicted proportion of examinees obtaining total score for DINA predictions; the smaller the difference between the lines, the better the model prediction. There is substantial misfit for total scores exceeding 80.

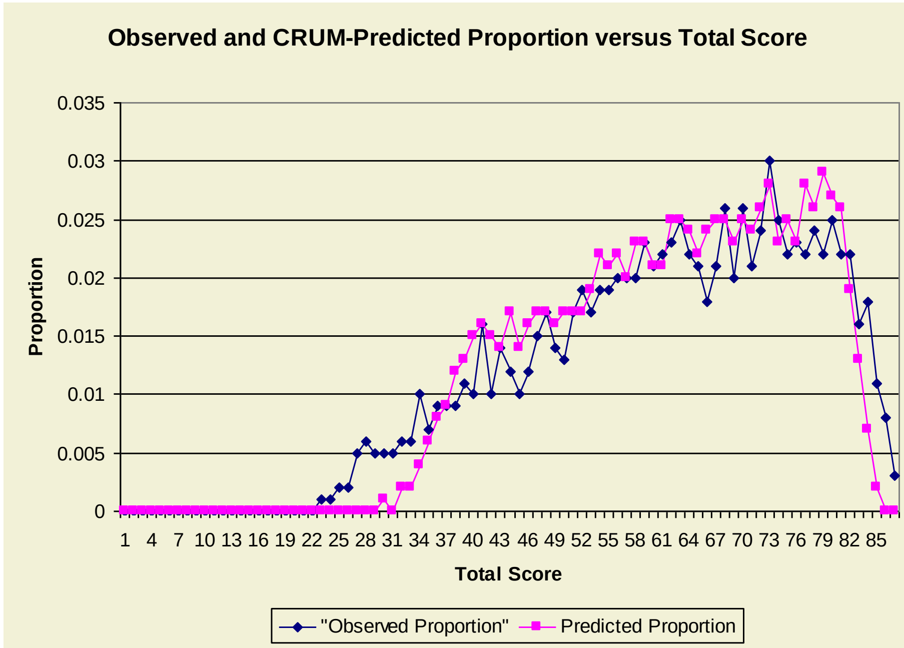

Figure 5. A plot of observed and predicted proportion of examinees obtaining total score for DINA predictions; the smaller the difference between the lines, the better the model prediction.

CHAPTER 4: DISCUSSION

Model Selection and Conclusion

Compared to the DINA, the CRUM appears to be less conservative, making it more likely for a given student to be credited with mastery of a given attribute. The CRUM also had more variability in its allocation of students among attribute mastery profiles; one could interpret this as greater discrimination between attribute profiles in the CRUM than in the DINA. Both the increased conservatism and the decreased variability in student distribution among classes in the DINA can be attributed to the non-compensatory nature of that model. In both the DINA and CRUM, a student is more likely to be classified as possessing an attribute if he correctly answers more items possessing that attribute, however the non-compensatory DINA, by looking only at interaction terms, is more conservative because of the inherent reliance on the presence of other attributes involved in the items.

A look at the relative fit indices in Table 7 for the two models indicates that the CRUM is a better fit. This is unsurprising, given the dramatic increase in the number of parameters estimated by the CRUM as opposed to the DINA. As neither the DINA nor the CRUM are nested within the other, a chi-squared test for improved model fit cannot be conducted; the AIC, however, does take into account the parsimony of the model and can be used for non-nested model comparison. Given the more refined distribution of attribute mastery patterns, and fewer cases of all-or-nothing attribute mastery probabilities within examinee profiles, it appears that the CRUM would be the more appropriate model for these items and cognitive attributes. Table 8 indicates greater item-level misfit for the DINA, as the MAD and RMSE are larger than those of the CRUM. Finally, when considering one measure of absolute model fit, a comparison of Figure 4 and Figure 5 indicates that the CRUM is far superior to the DINA in predicting proportions of students with a given total score, particularly in the range of T = (34,80). The DINA, on the other hand, consistently under-predicts the proportion for the majority of that same range (Figure 4), and then grossly over-predicts for the very top scores of T > 80.

A reliable means of assessing mastery of an individual attribute would be the inclusion of items that only possess said attribute; the current study, being a retrospective analysis of an test that was not constructed for the purpose of diagnosis of cognitive components, had only nine items that involved a single component, which was in each case either the Number and Computation blueprint standard or the Contextual Encoding cognitive component, which are the second and third attributes in the **Q-**matrix. One could expect those two attributes to be more precisely measured than the others because of their unique presence on those nine items. As none of the remaining eight attributes were uniquely represented on the test, their diagnosis is partially confounded with that of others for both models.

Limitations

Model estimation

MCMC methods are an alternative to ML methods that require derivatives for parameter estimation when dealing with a large number of attributes, as the state of current software programs does not allow for more than 2^9 latent classes using EM methods. Additionally, the appeal of the LCDM is the flexibility in model specification, potentially at a per-item basis. However the current study found that, even with the state-of-the-art LCDM estimation methods, there was limited customizability for model selection for each item: only pre-established models could be estimated, reducing the types of model comparison available. Additionally, more study needs to be done on the impact of one unstable parameter on the estimation of other parameters, both within and between items, as well as on the estimation of attribute mastery probabilities. One possible solution is to allow the user to thin out the parameter chain to reduce autocorrelation effects before calculating final parameter estimates, which is not automated with the current LCDM software.

***Q**-matrix specification*

The formulation of the LCDM and, therefore, any nested models, results in a forced classification of students as masters or non-masters of different attributes, when in fact some attributes could yield a more finely measured scaling of examinees, as in the case of traditional IRT. The cognitive attributes used for the present study contain more information than a simple dichotomy would reveal. For example, Contextual Encoding was originally scored as a count of the number of non-mathematical terms in an item; by recoding that as a binary variable, information about the nature of an item's Contextual Encoding was lost. Indeed, it can be argued that it is not the presence of the attributes on the items that one is interested in, but rather the extent to which the attributes are involved.

Item adequacy

The data used for the present study were students' responses on a live mathematics achievement test, which was designed from the traditional IRT perspective.

Thus, the phenomenon of having poorly represented cognitive attributes, as well as items where many attributes were present, contributed to the difficulty in obtaining a stable model and precision in parameter estimation, which in turn led to differences in diagnostic profiles for the examinees between the models. If an administration is truly interested in diagnosis, new tests designed with the attributes of interest in mind will need to be developed. The current study shows that current examinations can be used for diagnostic assessment, but better results would likely be obtained if the test were designed with the dual purpose of assessment and diagnosis in mind.

APPENDIX A

COGNITIVE COMPONENT DEFINITIONS AND ITEM SCORES

Table A 1. Cognitive Components and Their Definitions

Attribute	Definition
Translation	
Encoding	The total number of words and math terms, both explicit and implicit, in the stem and answer options. Equal to the sum of Mathematical and Context attributes.
Mathematical Encoding	The total number of mathematical terms, both implicit and explicit, in the stem and all answer options. This includes numerals, variables (e.g., *x*, *y*, *m*, etc.), axis labels, comparators (e.g., <, >, =), and implicit and explicit operators.
Contextual Encoding	The total number of words, excluding variables, in the stem and all answer options.
Translate Word Eqs	Indicator of whether the examinee needs to interpret an equation given in word (context) form.
Encode Diagram	Indicator of presence of a diagram, graph, or other figure, excluding tables, in the stem or answer options.
Reading Complexity	The maximum MS Word-determined reading level for stem and answer options.
Integration	
Generate Equations or Plausible Values	Indicator of whether examinee must generate or derive equations or possible values for variables in order to answer the item.
Recall Equations	Indicator of whether the examinee must recall known equations (e.g., formula for slope of a line, the Pythagorean theorem, etc.) in order to answer the item
Translate Diagram	Indicator for whether presented diagram or figure is necessary for problem solution.
Visualization	Indicator for whether examinee must draw or otherwise visualize a diagram or figure to understand or answer the item
Solution Planning	
Number of Subgoals	The total number of sub-steps necessary for answering an item (e.g., finding a slope for the equation of a line)

Table A1 (continued)

	Relative Definition of Variables	Indicator of whether one variable is defined only in terms of another.
Solution Execution		
	Procedural Knowledge	The maximum procedural knowledge necessary in solving the item

Table A1 (continued)

	1. Integers *	Indicator for whether ability of integers is necessary for solving the item
	2. Fractions *	Indicator for whether ability to manipulate fractions is necessary for solving the item
	3. Proportions *	Indicator for whether ability to manipulate proportions is necessary for solving the item
	4. Decimals *	Indicator for whether ability to manipulate decimals is necessary for solving the item
	5. Negative Numbers *	Indicator for whether ability to manipulate negative numbers is necessary for solving the item
	6. Exponents and Radicals *	Indicator for whether ability to evaluate squares or square roots is necessary for solving the item
	Number of Procedures	The total number of procedures necessary for solving the item (e.g., if two different fractions are involved in an equation to be solved Number of Procedures would be 2)
	Number of Computations	The total number of computations necessary for solving the item; including computations necessary in evaluating answer options and in stem
Decision Processing		
	Decision Confirmation Processing	Indicator for whether information found in distractors is necessary to eliminate options or answer item
	Bottom-up	Indicator for whether distractors aid in bottom-up solution of the problem
	Top-down	Indicator for whether distractors are differentiated in a top-down solution of the problem

* These attributes are used to explicate their subsuming components, and were not included in the analysis.

Table A 2. Item Profiles for Cognitive Components: Translation

Item	Encoding	Mathematical Encoding	Contextual Encoding	Translate Word Equations	Encode Diagram	Reading Complexity
1	41	21	20	1	0	5.6
2	39	10	29	1	0	2.7
3	58	5	53	1	0	5.5
4	68	34	34	0	1	7.4
5	71	29	42	0	0	11.9
6	51	14	37	1	0	6
7	20	0	20	1	0	9.9
8	26	17	9	0	0	3.6
9	42	4	38	1	0	7.4
10	38	1	37	1	0	9.6
11	32	24	8	1	0	4.9
12	45	3	42	1	0	8.2
13	47	8	39	0	0	6.4
14	29	15	14	0	0	1.4
15	33	9	24	0	0	8.2
16	73	16	57	0	0	9.1
17	86	12	74	0	0	7.7
18	91	16	75	0	0	8.1
19	50	17	33	0	1	5.8
20	44	16	28	0	1	4.2
21	42	14	28	1	0	4.8
22	47	20	27	0	0	6.7
23	61	16	45	0	1	8.1
24	87	34	53	1	0	9.7
25	91	34	57	1	0	6.1
26	170	73	97	1	1	9.1
27	117	44	73	1	0	8.2
28	86	38	48	1	0	8.8
29	130	61	69	0	0	6.9
30	79	38	41	1	0	11.1
31	64	15	49	0	0	4.5
32	28	20	8	0	0	3.7
33	71	28	43	1	0	8.8
34	34	26	8	0	0	2.8
35	69	14	55	0	0	6.7

Table A2 (continued)

36	117	1	116	0	0	6.8
37	61	13	48	0	1	7.3
38	63	19	44	0	1	12
39	42	8	34	0	0	10.5
40	49	10	39	0	0	9.5
41	42	9	33	0	1	6
42	44	8	36	0	0	7.7
43	34	7	27	0	0	8
44	54	17	37	0	1	4.7
45	67	8	59	0	1	7.1
46	65	10	55	0	1	9.3
47	71	10	61	0	1	5.3
48	33	6	27	0	0	4.8
49	33	6	27	0	0	3.6
50	58	9	49	0	1	9.1
51	66	32	34	1	0	7.6
52	36	6	30	0	0	6.7
53	32	5	27	0	0	10.7
54	55	9	46	0	0	6.8
55	70	18	52	0	0	6.9
56	27	10	17	0	0	3.6
57	39	12	27	0	1	6.8
58	56	19	37	0	0	7.8
59	59	27	32	0	0	8.1
60	82	32	50	0	0	8.9
61	70	33	37	0	0	7.6
62	61	31	30	0	0	6.9
63	65	22	43	0	1	4.7
64	86	31	55	1	0	5.4
65	36	26	10	0	0	7.6
66	12	7	5	0	0	7.3
67	34	4	30	0	0	5
68	25	18	7	0	0	3.9
69	12	7	5	0	0	0.6
70	22	15	7	0	0	4
71	20	12	8	0	0	6.7
72	20	0	20	0	0	0
73	17	1	16	0	0	6.2

Table A2 (continued)

74	26	6	20	0	0	9.9
75	54	39	15	0	1	1.5
76	39	0	39	0	0	6.9
77	32	15	17	0	1	5
78	74	14	60	0	0	5.3
79	50	38	12	0	1	0
80	44	30	14	0	1	0
81	44	34	10	0	1	0
82	46	26	20	0	1	4.8
83	32	14	18	1	0	7.1
84	97	41	56	1	0	9.3
85	110	30	80	1	0	7.4
86	89	54	35	1	0	7

Table A 3. Item Profiles for Cognitive Components: Integration

Item	Equation Given in Stem	Generate Equations or Plausible Values	Recall Equations	Recall Knowledge Principles	Translate Diagram	Visualization
1	0	0	0	1	0	0
2	0	0	0	0	0	0
3	0	1	0	0	0	0
4	1	0	0	0	0	0
5	1	0	0	0	0	1
6	0	1	0	0	0	0
7	0	0	0	1	0	0
8	1	0	0	1	0	0
9	0	1	0	1	0	0
10	0	0	0	1	0	0
11	0	0	0	1	0	0
12	0	1	0	0	0	0
13	0	0	1	1	0	0
14	0	0	1	1	0	0
15	0	0	1	1	0	0
16	0	0	1	1	0	0
17	0	0	0	1	0	0
18	0	0	1	1	0	0
19	1	1	0	1	1	0
20	1	0	0	1	1	0
21	1	0	0	1	0	1
22	1	0	0	1	0	0
23	1	0	0	1	0	0
24	0	1	0	1	0	0
25	0	1	0	0	0	0
26	1	0	0	0	1	1
27	0	1	0	0	0	0
28	0	1	0	0	0	0
29	1	0	0	0	0	0
30	0	1	0	0	0	0
31	1	0	0	0	0	0
32	1	0	0	0	0	0
33	0	1	0	1	0	1
34	1	0	0	0	0	0

Table A3 (continued)

35	1	0	0	0	0	0
36	0	0	0	1	0	0
37	0	0	0	1	0	0
38	0	0	0	1	1	0
39	0	0	0	1	0	0
40	0	0	0	1	0	0
41	0	0	0	1	1	0
42	0	0	0	1	0	0
43	0	0	0	0	0	1
44	0	0	0	0	0	0
45	0	0	0	0	1	0
46	0	0	0	0	0	0
47	0	0	0	0	0	0
48	0	0	0	0	0	0
49	0	0	0	0	0	0
50	0	0	1	0	0	0
51	0	0	0	0	0	0
52	0	0	1	0	0	1
53	0	0	1	0	0	1
54	0	0	0	1	0	0
55	0	0	0	1	0	0
56	0	0	0	1	0	1
57	0	0	0	1	1	0
58	0	1	0	0	0	0
59	0	1	0	0	0	0
60	0	1	0	0	0	0
61	0	1	0	0	0	0
62	0	1	0	0	0	0
63	0	1	0	0	1	0
64	0	1	0	0	0	0
65	1	0	0	0	0	0
66	1	0	0	0	0	0
67	0	1	0	0	0	0
68	1	0	0	0	0	0
69	1	0	0	0	0	0
70	1	0	0	0	0	0
71	0	0	0	1	0	0
72	0	0	0	1	0	0

Table A3 (continued)

73	0	0	0	1	0	0
74	0	0	0	1	0	0
75	0	0	1	0	1	0
76	0	0	0	1	0	0
77	0	0	0	1	1	0
78	1	0	0	1	0	1
79	0	0	1	0	1	0
80	0	0	1	0	1	0
81	0	0	1	0	1	0
82	0	0	0	1	1	0
83	0	1	0	0	0	0
84	0	1	0	0	0	1
85	1	0	0	0	0	0
86	0	1	0	0	0	0

Table A 4. Item Profiles for Cognitive Components: Solution Planning

Item	Number of Subgoals	Specify Subgoals	Relative Definition of Variables
1	0		0
2	0		0
3	0		0
4	0		0
5	0		0
6	0		0
7	0		0
8	0		0
9	0		0
10	0		0
11	0		0
12	0		0
13	4	calculate each of mean, median, mode, range	0
14	0		0
15	0		0
16	0		0
17	0		0
18	0		0
19	0		0
20	0		0
21	0		0
22	0		0
23	0		0
24	0		0
25	0		0
26	0		0
27	1	calculate slope	0
28	2	calculate slope, intercept	0
29	0		0
30	1	calculate slope	0
31	0		0
32	0		0
33	0		1
34	0		0

Table A4 (continued)

35	0		0
36	0		0
37	0		0
38	0		0
39	0		0
40	0		0
41	1	identify Pr[A]	0
42	0		0
43	0		0
44	0		0
45	0		0
46	0		0
47	0		0
48	1	Calculate 30% of orig. price before subtracting from orig. price	0
49	0		0
50	0		0
51	0		0
52	0		0
53	0		0
54	0		0
55	0		0
56	0		0
57	1	count contents of bags	0
58	0		0
59	0		0
60	0		0
61	0		0
62	0		0
63	0		0
64	0		0
65	0		0
66	0		0
67	0		0
68	0		0
69	0		0
70	0		0
71	0		0

Table A4 (continued)

72	0		0
73	0		0
74	0		0
75	0		0
76	0		0
77	0		0
78	0		1
79	0		0
80	0		0
81	0		0
82	0		0
83	0		0
84	1	calculate slope	0
85	1	calculate total number of glasses	0
86	0		0

Table A 5. Item Profiles for Cognitive Components: Solution Execution and Decision Processing

	Solution Execution			Decision Processing		
Item	Procedural Knowledge	Number of Procedures	Computations	Decision Confirmation Processing	Bottom-up	Top-down
1	0	0	0	1	0	1
2	2	2	1	0	0	0
3	0	0	0	0	0	0
4	0	0	0	1	0	1
5	0	0	0	1	0	1
6	2	1	0	1	0	1
7	0	0	1	0	0	0
8	1	1	3	0	0	0
9	0	0	0	1	1	0
10	4	1	0	1	0	1
11	2	2	2	0	0	0
12	0	0	0	1	1	0
13	1	1	17	1	0	1
14	0	0	2	0	0	0
15	5	2	6	0	0	0
16	4	2	4	0	0	0
17	0	0	0	1	0	1
18	1	1	12	1	0	1
19	6	2	7	1	1	0
20	6	2	4	0	0	0
21	6	2	4	0	0	0
22	6	2	17	1	1	0
23	6	2	4	0	0	0
24	1	1	2	1	1	0
25	1	1	1	1	0	1
26	0	0	2	1	1	0
27	1	1	3	1	0	1
28	1	1	4	1	0	1
29	5	2	8	1	1	0
30	1	1	4	1	0	1
31	1	1	2	0	0	0
32	5	3	2	0	0	0

Table A5 (continued)

33	0	0	0	1	0	1
34	1	1	2	0	0	0
35	4	2	2	0	0	0
36	0	0	0	1	1	0
37	2	1	1	0	0	0
38	0	1	2	1	0	1
39	2	1	2	0	0	0
40	2	1	2	0	0	0
41	2	1	2	0	0	0
42	2	1	2	0	0	0
43	2	1	3	0	0	0
44	2	2	3	0	0	0
45	2	1	3	0	0	0
46	4	2	3	0	0	0
47	2	1	3	0	0	0
48	4	3	4	0	0	0
49	4	3	3	0	0	0
50	6	2	2	0	0	0
51	0	0	0	1	0	1
52	0	0	3	0	0	0
53	4	1	2	0	0	0
54	4	2	4	0	0	0
55	2	2	2	0	0	0
56	2	1	2	0	0	0
57	2	1	5	0	0	0
58	0	0	0	1	0	1
59	0	0	0	1	0	1
60	0	0	0	1	0	1
61	0	0	1	1	0	1
62	0	0	0	1	0	1
63	0	0	1	1	0	1
64	0	0	0	1	0	1
65	0	0	0	1	0	1
66	5	1	1	0	0	0
67	0	0	0	1	1	0
68	6	2	7	0	0	0
69	5	1	1	0	0	0
70	0	0	5	0	0	0

Table A5 (continued)

71	5	0	4	1	1	0
72	0	0	0	1	0	1
73	0	0	0	1	1	0
74	0	0	0	1	1	0
75	5	1	1	0	0	0
76	0	0	0	1	1	0
77	0	0	0	1	1	0
78	0	0	2	1	1	0
79	5	2	3	0	0	0
80	5	1	1	0	0	0
81	2	1	2	0	0	0
82	4	1	0	1	1	0
83	0	0	0	1	0	1
84	0	0	0	1	0	1
85	0	0	1	1	1	0
86	2	0	0	1	0	1

APPENDIX B

CONVERGENCE RESULTS FOR FULL DINA AND CRUM MODEL ESTIMATION

Table B 1. ANOVA results for DINA slip parameters

Source	df	SS	MS	F
Model	1	2.21566	2.21566	52,468 *
Error	84	0.00355	0.00004	--
Total	85	2.21921*	--	--

* $p < 0.01$

Coefficient	Estimate	Std. Error	t-value
(Intercept)	0.007750	0.003692	2.099 *
s2.Full	0.992185	0.004332	229.060 *

* $p < 0.01$

Table B 2. ANOVA results for DINA guess parameters

Source	df	SS	MS	F
Model	1	3.3095	3.3095	66,388 *
Error	84	0.0042	0.0000	--
Total	85	3.3137	--	--

* $p < 0.01$

Coefficient	Estimate	Std. Error	t-value
(Intercept)	0.001962	0.002006	0.978
g2.Full	0.999008	0.003877	257.658 *

* $p < 0.01$

Table B 3. ANOVA results for CRUM intercepts

Source	df	SS	MS	F
Model	1	89.586	89.586	54,551*
Error	84	0.138	0.002	--
Total	85	89.724	--	--

* $p < 0.01$

Coefficient	Estimate	Std. Error	*t*-value
(Intercept)	-0.007741	0.004462	-1.735
xi.Full	1.002533	0.004292	233.562 *

* $p < 0.01$

Table B 4. ANOVA results for CRUM main effects

Source	df	SS	MS	F
Model	1	276.505	276.505	2,392.9*
Error	84	32.701	0.116	--
Total	85	309.206	--	–

* $p < 0.01$

Coefficient	Estimate	Std. Error	*t*-value
(Intercept)	-0.05498	0.03020	-1.821
xl.Full	1.07868	0.02205	48.918 *

* $p < 0.01$

Table B 5. Standardized differences for DINA parameters

Item	Slip	Guess
X1	0.0226	0.5136
X2	-0.0057	0.2960
X3	0.0126	0.4396
X4	0.0177	0.1262
X5	0.0020	0.2524
X6	-0.0029	-0.1976
X7	0.0292	0.6387
X8	0.0096	1.5653 *
X9	0.0008	-0.0453
X10	-0.0061	-0.2183
X11	0.0066	0.1505
X12	0.0073	0.3688
X13	0.0039	0.2792
X14	0.0260	0.3748
X15	0.0035	0.3426
X16	-0.0057	-0.2863
X17	-0.0001	-0.0651
X18	0.0012	-0.0816
X19	0.0007	0.0356
X20	0.0020	0.1849
X21	0.0021	0.5475
X22	0.0016	-0.1800
X23	-0.0021	-0.0460
X24	-0.0031	-0.2398
X25	0.0010	0.1372
X26	0.0003	-0.0391
X27	-0.0002	0.0742
X28	0.0006	0.2320
X29	-0.0008	-0.3097
X30	-0.0003	0.2884
X31	0.0048	-0.0625
X32	-0.0859	-0.0702
X33	-0.0003	0.0093
X34	-0.0527	-0.0428
X35	-0.0049	-0.6746
X36	-0.0189	-0.3343
X37	0.0004	-0.0123
X38	-0.0047	-0.1098

Table B5 (continued)

X39	0.0011	0.0348
X40	-0.0020	-0.1326
X41	-0.0015	0.0374
X42	-0.0008	-0.1299
X43	-0.0054	0.0265
X44	0.0010	0.0780
X45	-0.0017	-0.2843
X46	-0.0025	-0.0697
X47	-0.0016	-0.1835
X48	0.0007	0.1679
X49	-0.0003	0.0347
X50	0.0004	0.1775
X51	-0.0016	-0.1992
X52	0.0004	0.2174
X53	0.0018	0.3267
X54	-0.0015	-0.0686
X55	0.0018	0.0648
X56	-0.0087	-0.3880
X57	-0.0025	-0.2238
X58	0.0064	0.0745
X59	0.0051	0.1195
X60	-0.0019	0.0624
X61	0.0011	0.1448
X62	-0.0059	-0.0292
X63	-0.0008	-0.1716
X64	0.0035	0.2101
X65	0.0020	0.1170
X66	0.0092	0.8002
X67	-0.0062	-0.3567
X68	0.0029	0.2085
X69	-0.0007	0.1195
X70	0.0068	0.6689
X71	-0.0004	-0.2709
X72	0.0055	0.3726
X73	0.0166	0.1728
X74	-0.0078	-0.2873
X75	0.0070	0.0467
X76	-0.0011	-0.2199
X77	-0.0017	-0.3329
X78	-0.0016	-0.1381

Table B5 (continued)

X79	0.0013	0.1194
X80	0.0016	0.1801
X81	0.0043	0.2626
X82	-0.0065	-0.3555
X83	0.0030	0.5155
X84	0.0005	0.2034
X85	-0.0114	-0.3216
X86	0.0031	0.0690

* *parameter falls outside of 67% confidence limits*

Table B 6. Standardized differences for CRUM intercepts

Item	Standardized Difference
X1	-0.4248
X2	-0.4247
X3	0.2406
X4	-0.3779
X5	-0.3764
X6	-0.2101
X7	-0.1694
X8	-0.9422
X9	-0.5839
X10	-0.1029
X11	-0.8416
X12	0.0411
X13	-0.0740
X14	-0.5090
X15	-0.9917 *
X16	-0.3296
X17	0.0629
X18	-0.2878
X19	-0.1146
X20	-0.1324
X21	-0.9809 *
X22	-0.6837
X23	-0.0023
X24	0.3237
X25	0.2472
X26	0.1371
X27	0.3665
X28	-0.6180
X29	-0.1014
X30	0.3553
X31	0.3766
X32	0.0043
X33	0.0138
X34	0.1380
X35	0.7213
X36	0.7330
X37	0.0384

Table B6 (continued)

X38	-0.2223
X39	1.1172 *
X40	0.3652
X41	0.3119
X42	0.3877
X43	0.1835
X44	0.0084
X45	0.1749
X46	0.2750
X47	-0.2500
X48	0.3449
X49	-0.0042
X50	0.0258
X51	-0.0252
X52	0.6694
X53	-0.4217
X54	0.1569
X55	0.2249
X56	0.6585
X57	-0.1177
X58	0.1345
X59	0.1375
X60	-0.2540
X61	0.2485
X62	-0.2731
X63	-0.0845
X64	-0.2933
X65	-0.6662
X66	-0.7831
X67	0.0751
X68	-0.2917
X69	-0.8117
X70	-0.4535
X71	-0.3399
X72	-0.2797
X73	-0.1354
X74	-0.0211
X75	-0.0665
X76	0.7097
X77	-0.2070

Table B6 (continued)

X78	-0.2486
X79	-0.2780
X80	-0.1807
X81	-0.1642
X82	-0.1960
X83	0.0630
X84	-0.3944
X85	0.1295
X86	0.1825

* *Parameter falls outside the 67% confidence limits*

Table B 7. Standardized differences for CRUM main effects parameters

Parameter	Standardized Difference
X1.2	0.0077
X2.2	0.3161
X3.6	-0.9062
X3.4	0.4571
X3.3	0.2371
X3.2	-0.1795
X4.5	-0.4398
X4.3	-0.4240
X4.2	-0.2204
X5.7	0.1415
X5.4	0.2485
X5.3	-0.3187
X5.2	-0.3288
X6.6	0.2981
X6.4	-0.0448
X6.3	0.1311
X6.2	-0.5971
X7.2	0.4437
X8.2	-0.1219
X9.10	0.0931
X9.6	0.4995
X9.4	-0.6133
X9.2	-0.4445
X10.4	-0.2562
X10.2	0.0223
X10.1	0.0836
X11.2	-0.2653
X11.1	0.0393
X12.10	0.0441
X12.6	0.3920
X12.4	-0.2806
X12.2	0.0317
X12.1	-0.2893
X13.8	-0.4726
X13.4	0.1620
X13.2	-0.0106
X13.1	0.2447

Table B7 (continued)

X14.2	0.3714
X14.1	-0.5331
X15.2	-0.2195
X15.1	-0.0993
X16.4	0.2891
X16.3	0.0147
X16.2	-0.4540
X16.1	0.1906
X17.4	0.1476
X17.3	0.1450
X17.2	-0.1676
X17.1	-0.2158
X18.4	0.6913
X18.3	0.0089
X18.2	-0.5255
X18.1	0.1144
X19.10	-0.4098
X19.6	-0.7223
X19.5	0.3222
X19.3	0.4560
X19.1	0.3041
X20.5	0.1300
X20.1	0.0682
X21.7	-0.8339
X21.1	0.2550
X22.10	-0.9698
X22.1	0.1358
X23.5	0.0924
X23.4	-0.2164
X23.3	0.3672
X23.1	0.0744
X24.10	-0.3772
X24.6	0.1924
X24.4	0.5292
X24.3	-0.3031
X24.1	-0.1967
X25.6	-0.1805
X25.4	0.5970
X25.3	-0.1396
X25.1	0.2981

Table B7 (continued)

X26.10	-0.9337
X26.7	0.2265
X26.5	0.0487
X26.4	0.7219
X26.3	0.1330
X26.1	0.3079
X27.8	0.2849
X27.6	0.0772
X27.4	0.4698
X27.3	0.0065
X27.1	-0.4687
X28.8	0.1523
X28.6	0.3085
X28.4	0.1741
X28.3	-0.3282
X28.1	-0.2477
X29.10	0.0661
X29.4	0.3172
X29.3	-0.0844
X29.1	0.0025
X30.8	-0.5269
X30.6	-0.1895
X30.4	0.7857
X30.3	0.5039
X30.1	0.2247
X31.4	0.4430
X31.3	-0.3297
X33.9	-0.7446
X33.7	-0.3502
X33.6	0.8361
X33.4	-0.2031
X33.3	0.4675
X35.4	0.5238
X35.3	0.1542
X36.1	0.8603
X36.4	-0.4353
X36.3	-0.2205
X37.5	-0.3267
X37.4	-0.1895
X37.3	-0.2764

Table B7 (continued)

X38.5	0.8914
X38.4	-1.1200 *
X38.3	0.2451
X39.4	-0.2388
X40.4	-0.0132
X40.3	-0.3190
X40.1	0.0881
X41.8	0.1948
X41.5	-0.0340
X41.4	-0.3864
X41.1	-0.1813
X42.4	-0.0245
X42.3	0.0021
X42.1	-0.1368
X43.7	-0.5178
X43.4	0.2313
X43.1	0.1728
X44.5	0.0410
X44.4	-0.6672
X44.3	0.5198
X44.1	0.0974
X45.5	-0.3119
X45.4	-0.3279
X45.3	0.5445
X45.1	-0.0003
X46.5	0.2053
X46.4	-0.5628
X46.3	0.3704
X46.1	-0.0639
X47.5	-0.5797
X47.4	-0.3715
X47.3	0.6779
X47.1	0.4687
X48.8	1.0498 *
X48.4	0.2488
X48.2	-0.7882
X48.1	-0.5312
X49.4	0.2728
X49.2	-0.1424
X49.1	-0.2509

Table B7 (continued)

X50.5	0.4960
X50.4	-1.1041 *
X50.3	0.2689
X50.2	0.7091
X50.1	0.0915
X51.4	0.3422
X51.3	1.0845 *
X51.2	-0.1754
X51.1	0.6500
X52.7	-0.1693
X52.4	0.5701
X52.2	-0.0231
X52.1	-0.2058
X53.7	0.1564
X53.4	0.0855
X53.2	-0.5962
X53.1	0.2556
X54.4	0.1080
X54.3	0.1908
X54.1	0.0506
X55.4	0.4425
X55.3	-0.3354
X55.1	0.4844
X56.7	1.0477 *
X56.1	-0.1501
X57.8	0.1636
X57.5	-0.3702
X57.4	0.1635
X57.1	0.3104
X58.6	-0.1108
X58.4	0.3026
X58.3	-0.1371
X58.2	-0.2623
X59.6	0.1358
X59.4	0.0850
X59.3	-0.0826
X59.2	-0.0794
X60.6	-0.2073
X60.4	0.1064
X60.3	-0.0848

Table B7 (continued)

X60.2	-0.1719
X61.6	-0.0790
X61.4	-0.0360
X61.3	0.1979
X61.2	-0.1206
X62.6	-0.3819
X62.4	0.1057
X62.3	0.0345
X62.2	-0.1521
X63.6	-0.2610
X63.5	-0.0475
X63.4	-0.3545
X63.3	-0.0278
X63.2	0.1688
X64.6	0.3302
X64.4	-0.4126
X64.3	-0.2422
X64.2	0.0208
X65.3	-0.0006
X65.2	-0.2674
X66.2	-0.1105
X67.10	-0.1639
X67.6	-0.6730
X67.4	-0.1797
X67.3	0.1469
X67.2	0.0074
X67.1	0.5397
X68.2	0.4041
X68.1	-0.2345
X69.2	-0.0259
X69.1	0.0814
X70.2	0.1067
X70.1	-0.0120
X71.10	-0.1463
X71.2	0.2341
X71.1	0.0425
X72.4	-0.2329
X72.2	0.2986
X72.1	-0.3499
X73.10	-0.8016

Table B7 (continued)

X73.2	0.7923
X73.1	0.0660
X74.10	0.6016
X74.4	0.0831
X74.2	-0.6733
X74.1	-0.0795
X75.5	0.5678
X75.3	-0.1035
X75.1	-0.6137
X76.10	0.1191
X76.4	0.2208
X76.1	0.0204
X77.10	0.4683
X77.5	-0.3624
X77.1	-0.3833
X78.10	0.1042
X78.9	2.8328 *
X78.7	-0.1539
X78.4	-0.1769
X78.3	-0.5806
X78.1	-0.6827
X79.5	-0.1856
X79.3	0.2354
X79.1	0.0404
X80.5	0.0752
X80.1	-0.2495
X81.5	0.0452
X81.1	-0.0012
X82.10	0.2519
X82.5	-0.3397
X82.1	0.1134
X83.6	1.1073 *
X83.1	-0.6136
X84.8	0.2892
X84.7	0.3435
X84.6	-0.0695
X84.4	-0.6449
X84.3	-0.4631
X84.1	0.0373
X85.1	-0.1076

Table B7 (continued)

X85.8	-0.4445
X85.4	0.0112
X85.3	0.3223
X85.1.1	0.4664
X86.6	0.0070
X86.4	-0.0612
X86.3	0.1794
X86.1	0.0300

** Parameter falls outside the 67% confidence limits*

Table B 8. Probabilities of successful item completion for items flagged parameters

Flagged Parameter	Model	Pr[X=1] First	Pr[X=1] Last	Pr[X=1] Overall	Difference
X8	DINA	0.49173	0.4481948	0.4706412	0.04353515
X15	CRUM	0.941388	0.946523	0.943928	-0.005135
X21	CRUM	0.996533	0.996809	0.996656	-0.000275
X38.4	CRUM	0.838483	0.857394	0.846717	-0.018911
X39	CRUM	0.917652	0.908185	0.913717	0.009467
X48.8	CRUM	0.989620	0.987694	0.988930	0.001926
X50.4	CRUM	0.999101	0.999308	0.999233	-0.000207
X51.3	CRUM	0.999999	0.999994	0.999998	0.000006
X56.7	CRUM	0.983164	0.981206	0.982196	0.001958
X78.9	CRUM	0.999940	0.993231	0.999337	0.006709
X83.6	CRUM	0.999982	0.999934	0.999965	0.000049

Table B 9. Flagged items and attributes and their properties

Item	Model	Flagged Attribute	Facility	Item's Attribute Saturation	Attribute's Test Representation
8	DINA	guess	0.878	1	41
15	CRUM	intercept	0.812	2	N/A
21	CRUM	intercept	0.711	2	N/A
38	CRUM	4	0.541	3	57
39	CRUM	intercept	0.853	1	N/A
48	CRUM	8	0.499	4	9
50	CRUM	4	0.336	5	57
51	CRUM	3	0.945	4	47
56	CRUM	7	0.546	2	10
78	CRUM	9	0.613	6	2
83	CRUM	6	0.817	2	22

Table B 10. Parameter summary for DINA slip parameters across Geweke windows (p1 = p2 = 0.40)

Item	First Mean	Last Mean	Diff	First Variance	Last Variance
X1	0.5661	0.5432	0.0229	0.0056	0.0051
X2	0.8223	0.8282	-0.0058	0.0008	0.0010
X3	0.8211	0.8115	0.0096	0.0004	0.0004
X4	0.3823	0.3810	0.0014	0.0001	0.0001
X5	0.9146	0.9122	0.0024	0.0003	0.0003
X6	0.9470	0.9501	-0.0031	0.0000	0.0000
X7	0.6483	0.6187	0.0296	0.0047	0.0040
X8	0.9610	0.9490	0.0120	0.0001	0.0001
X9	0.9708	0.9699	0.0009	0.0000	0.0000
X10	0.8702	0.8753	-0.0051	0.0003	0.0003
X11	0.7801	0.7736	0.0064	0.0008	0.0010
X12	0.9224	0.9165	0.0059	0.0001	0.0001
X13	0.8970	0.8934	0.0036	0.0001	0.0002
X14	0.6180	0.5917	0.0263	0.0030	0.0032
X15	0.8805	0.8768	0.0037	0.0003	0.0003
X16	0.8678	0.8731	-0.0052	0.0002	0.0002
X17	0.7949	0.7950	-0.0001	0.0007	0.0006
X18	0.8761	0.8747	0.0014	0.0004	0.0004
X19	0.9501	0.9493	0.0008	0.0001	0.0001
X20	0.8883	0.8862	0.0020	0.0002	0.0002
X21	0.9670	0.9644	0.0027	0.0000	0.0000
X22	0.9686	0.9666	0.0020	0.0000	0.0000
X23	0.9160	0.9184	-0.0025	0.0002	0.0002
X24	0.8100	0.8133	-0.0033	0.0007	0.0007
X25	0.9074	0.9063	0.0011	0.0001	0.0002
X26	0.9560	0.9557	0.0003	0.0000	0.0000
X27	0.8084	0.8086	-0.0002	0.0005	0.0006
X28	0.9383	0.9377	0.0006	0.0000	0.0001
X29	0.9119	0.9128	-0.0009	0.0002	0.0002
X30	0.9572	0.9575	-0.0003	0.0000	0.0000
X31	0.7238	0.7190	0.0048	0.0015	0.0015
X32	0.2330	0.2336	-0.0006	0.0001	0.0001
X33	0.8824	0.8827	-0.0003	0.0002	0.0002
X34	0.0692	0.0695	-0.0002	0.0000	0.0000
X35	0.9435	0.9491	-0.0057	0.0001	0.0001
X36	0.6152	0.6296	-0.0143	0.0010	0.0011
X37	0.8569	0.8563	0.0006	0.0030	0.0029
X38	0.7879	0.7912	-0.0033	0.0004	0.0004
X39	0.9364	0.9350	0.0014	0.0002	0.0002
X40	0.8909	0.8932	-0.0023	0.0003	0.0003
X41	0.8975	0.8990	-0.0015	0.0001	0.0001
X42	0.8769	0.8779	-0.0010	0.0004	0.0005
X43	0.7095	0.7147	-0.0052	0.0014	0.0015

Table B10 (continued)

X44	0.9652	0.9640	0.0011	0.0000	0.0000
X45	0.9572	0.9591	-0.0019	0.0000	0.0000
X46	0.8992	0.9021	-0.0029	0.0003	0.0002
X47	0.9167	0.9182	-0.0016	0.0001	0.0001
X48	0.9440	0.9433	0.0007	0.0000	0.0001
X49	0.8876	0.8880	-0.0004	0.0004	0.0004
X50	0.9903	0.9898	0.0005	0.0000	0.0000
X51	0.9128	0.9151	-0.0023	0.0018	0.0022
X52	0.9712	0.9707	0.0005	0.0000	0.0000
X53	0.9683	0.9663	0.0020	0.0000	0.0000
X54	0.9222	0.9233	-0.0011	0.0001	0.0001
X55	0.9435	0.9419	0.0017	0.0001	0.0001
X56	0.9035	0.9113	-0.0078	0.0001	0.0001
X57	0.9546	0.9570	-0.0024	0.0000	0.0000
X58	0.5775	0.5708	0.0068	0.0046	0.0042
X59	0.5093	0.5048	0.0045	0.0024	0.0021
X60	0.6610	0.6628	-0.0019	0.0020	0.0021
X61	0.8542	0.8531	0.0011	0.0003	0.0003
X62	0.6229	0.6280	-0.0051	0.0015	0.0014
X63	0.9743	0.9751	-0.0008	0.0000	0.0000
X64	0.9013	0.8971	0.0042	0.0003	0.0004
X65	0.6569	0.6549	0.0020	0.0027	0.0034
X66	0.9202	0.9110	0.0092	0.0003	0.0002
X67	0.8992	0.9049	-0.0057	0.0002	0.0001
X68	0.9078	0.9051	0.0028	0.0002	0.0002
X69	0.9324	0.9331	-0.0007	0.0001	0.0001
X70	0.9465	0.9395	0.0070	0.0001	0.0001
X71	0.9244	0.9249	-0.0005	0.0001	0.0001
X72	0.9015	0.8963	0.0052	0.0001	0.0002
X73	0.6895	0.6758	0.0137	0.0010	0.0009
X74	0.7261	0.7321	-0.0061	0.0006	0.0008
X75	0.8148	0.8061	0.0087	0.0020	0.0025
X76	0.7928	0.7939	-0.0011	0.0006	0.0007
X77	0.7804	0.7823	-0.0019	0.0016	0.0017
X78	0.9296	0.9313	-0.0017	0.0001	0.0001
X79	0.9472	0.9458	0.0014	0.0001	0.0001
X80	0.7987	0.7970	0.0017	0.0012	0.0010
X81	0.9022	0.8978	0.0045	0.0001	0.0002
X82	0.8479	0.8543	-0.0063	0.0004	0.0002
X83	0.9705	0.9663	0.0041	0.0001	0.0001
X84	0.9271	0.9266	0.0005	0.0001	0.0001
X85	0.7956	0.8064	-0.0108	0.0006	0.0005
X86	0.8855	0.8835	0.0019	0.0002	0.0003

Table B 11. Parameter summary for DINA guess parameters across Geweke windows (p1 = p2 = 0.40)

Item	First Mean	Last Mean	Diff	First Variance	Last Variance
X1	0.1453	0.1353	0.0100	0.0003	0.0002
X2	0.3808	0.3729	0.0079	0.0004	0.0006
X3	0.5879	0.5802	0.0077	0.0002	0.0002
X4	0.3797	0.3782	0.0015	0.0001	0.0001
X5	0.3799	0.3753	0.0046	0.0002	0.0002
X6	0.6487	0.6519	-0.0033	0.0002	0.0002
X7	0.1904	0.1759	0.0145	0.0004	0.0003
X8	0.5146	0.4691	0.0456	0.0006	0.0006
X9	0.7155	0.7161	-0.0007	0.0002	0.0001
X10	0.6281	0.6322	-0.0042	0.0002	0.0002
X11	0.3486	0.3456	0.0030	0.0003	0.0003
X12	0.7597	0.7546	0.0052	0.0001	0.0001
X13	0.6275	0.6233	0.0042	0.0001	0.0002
X14	0.1697	0.1642	0.0056	0.0001	0.0002
X15	0.4331	0.4266	0.0066	0.0002	0.0002
X16	0.5434	0.5484	-0.0051	0.0002	0.0002
X17	0.3672	0.3683	-0.0011	0.0002	0.0002
X18	0.3358	0.3372	-0.0014	0.0002	0.0002
X19	0.5887	0.5881	0.0005	0.0002	0.0002
X20	0.4888	0.4858	0.0031	0.0002	0.0002
X21	0.5748	0.5642	0.0106	0.0003	0.0002
X22	0.6004	0.6035	-0.0031	0.0002	0.0002
X23	0.3977	0.3984	-0.0007	0.0002	0.0002
X24	0.3160	0.3194	-0.0034	0.0001	0.0002
X25	0.5521	0.5499	0.0023	0.0002	0.0002
X26	0.7640	0.7646	-0.0005	0.0001	0.0001
X27	0.3851	0.3840	0.0011	0.0002	0.0001
X28	0.6408	0.6376	0.0032	0.0001	0.0001
X29	0.4476	0.4529	-0.0054	0.0002	0.0002
X30	0.5912	0.5869	0.0044	0.0002	0.0001
X31	0.2686	0.2697	-0.0011	0.0002	0.0002
X32	0.2330	0.2336	-0.0006	0.0001	0.0001
X33	0.5841	0.5840	0.0001	0.0001	0.0002
X34	0.0692	0.0695	-0.0002	0.0000	0.0000
X35	0.5579	0.5727	-0.0148	0.0003	0.0004
X36	0.3420	0.3473	-0.0053	0.0001	0.0002
X37	0.1503	0.1504	-0.0002	0.0001	0.0001
X38	0.5856	0.5874	-0.0018	0.0002	0.0002
X39	0.4054	0.4046	0.0008	0.0004	0.0004
X40	0.3570	0.3592	-0.0022	0.0002	0.0002
X41	0.5300	0.5295	0.0005	0.0001	0.0001
X42	0.3051	0.3074	-0.0023	0.0002	0.0002
X43	0.2689	0.2685	0.0004	0.0001	0.0002

Table B11 (continued)

X44	0.6346	0.6333	0.0013	0.0002	0.0002
X45	0.6288	0.6334	-0.0046	0.0002	0.0002
X46	0.3806	0.3816	-0.0011	0.0002	0.0001
X47	0.6383	0.6409	-0.0026	0.0001	0.0002
X48	0.7195	0.7170	0.0025	0.0001	0.0001
X49	0.3564	0.3558	0.0006	0.0002	0.0002
X50	0.8988	0.8971	0.0018	0.0001	0.0001
X51	0.1320	0.1343	-0.0023	0.0001	0.0001
X52	0.6998	0.6966	0.0031	0.0001	0.0001
X53	0.7537	0.7487	0.0050	0.0002	0.0001
X54	0.8026	0.8037	-0.0010	0.0002	0.0002
X55	0.7692	0.7681	0.0011	0.0002	0.0002
X56	0.6605	0.6664	-0.0059	0.0002	0.0002
X57	0.7753	0.7783	-0.0030	0.0001	0.0001
X58	0.1286	0.1277	0.0009	0.0001	0.0001
X59	0.1750	0.1734	0.0015	0.0001	0.0001
X60	0.2120	0.2111	0.0009	0.0001	0.0002
X61	0.4713	0.4690	0.0023	0.0002	0.0002
X62	0.2728	0.2732	-0.0004	0.0002	0.0001
X63	0.8479	0.8497	-0.0018	0.0001	0.0000
X64	0.3389	0.3356	0.0033	0.0002	0.0002
X65	0.2160	0.2135	0.0025	0.0003	0.0003
X66	0.6095	0.5873	0.0222	0.0006	0.0004
X67	0.6231	0.6289	-0.0058	0.0002	0.0002
X68	0.6096	0.6050	0.0046	0.0003	0.0003
X69	0.6189	0.6165	0.0024	0.0003	0.0003
X70	0.6788	0.6653	0.0134	0.0003	0.0002
X71	0.5644	0.5693	-0.0048	0.0002	0.0002
X72	0.5975	0.5912	0.0063	0.0002	0.0002
X73	0.3591	0.3562	0.0029	0.0002	0.0002
X74	0.4438	0.4485	-0.0047	0.0002	0.0002
X75	0.1691	0.1685	0.0006	0.0001	0.0001
X76	0.3777	0.3814	-0.0036	0.0002	0.0002
X77	0.2140	0.2186	-0.0046	0.0001	0.0001
X78	0.5807	0.5828	-0.0021	0.0001	0.0002
X79	0.6232	0.6212	0.0020	0.0002	0.0002
X80	0.2887	0.2858	0.0029	0.0002	0.0001
X81	0.5125	0.5083	0.0042	0.0002	0.0002
X82	0.4568	0.4623	-0.0055	0.0002	0.0001
X83	0.4064	0.3975	0.0089	0.0002	0.0002
X84	0.6883	0.6855	0.0028	0.0001	0.0002
X85	0.4010	0.4059	-0.0050	0.0002	0.0001
X86	0.7889	0.7880	0.0009	0.0001	0.0001

Table B 12. Parameter summary for CRUM parameters across Geweke windows (p1 = p2 = 0.40)

Item	Parameter	First Mean	Last Mean	Difference	First Variance	Last Variance
X1	intercept	-2.3159	-2.2565	-0.0594	0.0127	0.0136
X1	2	1.7291	1.7270	0.0021	0.0539	0.0445
X10	intercept	0.8595	0.8695	-0.0100	0.0060	0.0069
X10	1	0.9693	0.9565	0.0128	0.0151	0.0165
X10	2	0.7107	0.7077	0.0030	0.0102	0.0149
X10	4	0.7071	0.7434	-0.0363	0.0100	0.0201
X11	intercept	-0.6867	-0.6136	-0.0731	0.0048	0.0054
X11	1	0.8609	0.8515	0.0094	0.0370	0.0417
X11	2	1.5333	1.5756	-0.0424	0.0174	0.0161
X12	intercept	1.2557	1.2517	0.0040	0.0073	0.0044
X12	1	0.6961	0.7390	-0.0428	0.0163	0.0112
X12	2	0.1066	0.1034	0.0032	0.0072	0.0054
X12	4	0.2591	0.3006	-0.0415	0.0160	0.0119
X12	6	0.7665	0.7010	0.0655	0.0214	0.0131
X12	10	0.1517	0.1463	0.0054	0.0088	0.0119
X13	intercept	0.8147	0.8221	-0.0074	0.0070	0.0060
X13	1	1.0443	1.0042	0.0402	0.0186	0.0167
X13	2	0.6958	0.6975	-0.0018	0.0190	0.0170
X13	4	0.3228	0.2973	0.0255	0.0152	0.0192
X13	8	0.5399	0.6233	-0.0834	0.0205	0.0214
X14	intercept	-1.5555	-1.4969	-0.0586	0.0081	0.0104
X14	1	1.0048	1.2597	-0.2548	0.1202	0.2165
X14	2	2.0062	1.9113	0.0949	0.0435	0.0435
X15	intercept	-0.2739	-0.1768	-0.0971	0.0062	0.0068
X15	1	0.9708	0.9964	-0.0256	0.0422	0.0481
X15	2	2.0499	2.0871	-0.0373	0.0195	0.0186
X16	intercept	0.5429	0.5745	-0.0316	0.0060	0.0064
X16	1	0.7252	0.6928	0.0324	0.0199	0.0180
X16	2	0.6072	0.6859	-0.0787	0.0211	0.0179
X16	3	0.7161	0.7139	0.0022	0.0162	0.0134
X16	4	0.6821	0.6368	0.0453	0.0166	0.0159
X17	intercept	0.0108	0.0047	0.0061	0.0051	0.0085
X17	1	0.7935	0.8468	-0.0534	0.0373	0.0479
X17	2	0.9479	0.9778	-0.0299	0.0205	0.0224
X17	3	0.3774	0.3484	0.0290	0.0299	0.0202
X17	4	1.1185	1.0916	0.0269	0.0217	0.0228
X18	intercept	-0.0156	0.0136	-0.0292	0.0063	0.0079

Table B12 (continued)

X18	1	0.6573	0.6184	0.0389	0.0791	0.0729
X18	2	1.4793	1.6127	-0.1334	0.0478	0.0333
X18	3	0.9796	0.9771	0.0025	0.0521	0.0529
X18	4	1.0879	0.9469	0.1410	0.0307	0.0218
X19	intercept	1.0672	1.0781	-0.0109	0.0058	0.0064
X19	1	0.9567	0.8833	0.0734	0.0405	0.0354
X19	3	0.3068	0.2296	0.0773	0.0195	0.0185
X19	5	1.3238	1.2479	0.0760	0.0334	0.0444
X19	6	0.4310	0.5931	-0.1621	0.0328	0.0351
X19	10	0.9129	1.0186	-0.1057	0.0365	0.0600
X2	intercept	-1.0224	-0.9833	-0.0391	0.0051	0.0066
X2	2	1.7963	1.7454	0.0509	0.0168	0.0183
X20	intercept	0.3730	0.3836	-0.0106	0.0033	0.0061
X20	1	0.3095	0.2967	0.0128	0.0245	0.0216
X20	5	2.6398	2.6193	0.0205	0.0165	0.0169
X21	intercept	0.8326	0.9156	-0.0829	0.0039	0.0066
X21	1	0.8851	0.8071	0.0779	0.0603	0.0662
X21	7	3.9057	4.0891	-0.1833	0.0241	0.0484
X22	intercept	0.9496	1.0321	-0.0825	0.0102	0.0087
X22	1	1.1191	1.0774	0.0417	0.0712	0.0463
X22	10	3.4847	3.7597	-0.2750	0.0476	0.0656
X23	intercept	0.2746	0.2748	-0.0002	0.0067	0.0081
X23	1	0.2883	0.2714	0.0168	0.0339	0.0345
X23	3	0.3832	0.3038	0.0794	0.0347	0.0241
X23	4	0.3219	0.3609	-0.0390	0.0220	0.0209
X23	5	3.2363	3.2115	0.0248	0.0465	0.0513
X24	intercept	-0.1532	-0.1858	0.0326	0.0069	0.0065
X24	1	0.3860	0.4378	-0.0518	0.0455	0.0480
X24	3	0.5122	0.5819	-0.0697	0.0379	0.0300
X24	4	1.2286	1.1276	0.1011	0.0234	0.0261
X24	6	0.9629	0.8848	0.0782	0.1132	0.1039
X24	10	0.4435	0.5241	-0.0806	0.0313	0.0288
X25	intercept	0.8199	0.7955	0.0244	0.0068	0.0059
X25	1	0.6016	0.5392	0.0624	0.0315	0.0245
X25	3	2.0618	2.0898	-0.0280	0.0298	0.0207
X25	4	0.8501	0.7483	0.1018	0.0219	0.0143
X25	6	0.2471	0.2819	-0.0348	0.0237	0.0270
X26	intercept	1.9510	1.9352	0.0158	0.0094	0.0075
X26	1	1.0522	1.0081	0.0441	0.0147	0.0117
X26	3	1.3243	1.3020	0.0223	0.0181	0.0199

Table B12 (continued)

X26	4	0.5898	0.4687	0.1210	0.0212	0.0138
X26	5	0.0945	0.0898	0.0048	0.0060	0.0071
X26	7	0.1977	0.1662	0.0315	0.0124	0.0139
X26	10	0.1377	0.2881	-0.1504	0.0096	0.0327
X27	intercept	0.0636	0.0314	0.0322	0.0049	0.0056
X27	1	0.4557	0.5657	-0.1101	0.0331	0.0442
X27	3	1.4167	1.4154	0.0013	0.0277	0.0260
X27	4	0.7732	0.6986	0.0746	0.0173	0.0158
X27	6	0.1680	0.1553	0.0128	0.0196	0.0157
X27	8	0.2376	0.1815	0.0561	0.0271	0.0233
X28	intercept	1.0664	1.1271	-0.0607	0.0064	0.0064
X28	1	0.6451	0.6926	-0.0476	0.0266	0.0205
X28	3	2.0527	2.1112	-0.0585	0.0193	0.0249
X28	4	0.2782	0.2529	0.0253	0.0141	0.0140
X28	6	0.2257	0.1703	0.0555	0.0245	0.0158
X28	8	0.1968	0.1705	0.0262	0.0199	0.0194
X29	intercept	0.5714	0.5814	-0.0100	0.0059	0.0075
X29	1	0.9005	0.8998	0.0007	0.0519	0.0439
X29	3	2.4014	2.4240	-0.0225	0.0522	0.0380
X29	4	1.0579	1.0060	0.0519	0.0169	0.0198
X29	10	0.3538	0.3423	0.0115	0.0217	0.0175
X3	intercept	0.4332	0.4086	0.0246	0.0076	0.0057
X3	2	0.1202	0.1400	-0.0199	0.0069	0.0106
X3	3	0.1709	0.1408	0.0301	0.0118	0.0087
X3	4	0.3865	0.3191	0.0674	0.0159	0.0116
X3	6	1.0182	1.1792	-0.1610	0.0243	0.0146
X30	intercept	1.0493	1.0144	0.0349	0.0071	0.0051
X30	1	0.9370	0.8800	0.0571	0.0451	0.0388
X30	3	2.4288	2.3388	0.0900	0.0201	0.0235
X30	4	0.6802	0.5508	0.1294	0.0193	0.0158
X30	6	0.2383	0.2798	-0.0415	0.0295	0.0371
X30	8	0.1836	0.2923	-0.1087	0.0206	0.0440
X31	intercept	-0.8800	-0.9178	0.0378	0.0064	0.0073
X31	3	0.9570	1.0293	-0.0723	0.0339	0.0283
X31	4	1.4022	1.3131	0.0891	0.0283	0.0242
X32	intercept	-1.1880	-1.1883	0.0002	0.0023	0.0017
X33	intercept	0.6866	0.6854	0.0012	0.0046	0.0054
X33	3	0.2287	0.1657	0.0630	0.0128	0.0107
X33	4	0.8612	0.8878	-0.0266	0.0088	0.0166
X33	6	0.3242	0.1724	0.1518	0.0258	0.0143

Table B12 (continued)

X33	7	0.1102	0.1467	-0.0365	0.0063	0.0090
X33	9	2.4476	3.2329	-0.7853	0.9400	0.3444
X34	intercept	-2.5972	-2.6094	0.0122	0.0052	0.0053
X35	intercept	0.5498	0.4810	0.0688	0.0066	0.0049
X35	3	2.0014	1.9717	0.0297	0.0227	0.0288
X35	4	1.7056	1.6263	0.0792	0.0147	0.0163
X36	intercept	-0.2735	-0.3441	0.0706	0.0063	0.0059
X36	3	0.0625	0.0780	-0.0156	0.0031	0.0038
X36	4	0.8720	0.9541	-0.0821	0.0277	0.0157
X36	10	0.6174	0.4542	0.1631	0.0290	0.0138
X37	intercept	-0.9987	-1.0032	0.0045	0.0102	0.0075
X37	3	0.6526	0.7624	-0.1098	0.0955	0.1248
X37	4	2.5216	2.6033	-0.0817	0.1029	0.1660
X37	5	1.2484	1.3764	-0.1279	0.0854	0.1357
X38	intercept	0.5108	0.5297	-0.0189	0.0043	0.0058
X38	3	0.6461	0.6156	0.0305	0.0088	0.0133
X38	4	0.3187	0.4655	-0.1468	0.0099	0.0145
X38	5	0.2253	0.1254	0.0999	0.0095	0.0062
X39	intercept	-0.4708	-0.5900	0.1192	0.0067	0.0093
X39	4	2.8544	2.9058	-0.0514	0.0309	0.0308
X4	intercept	-0.4811	-0.4643	-0.0169	0.0013	0.0014
X4	2	0.0114	0.0153	-0.0039	0.0003	0.0001
X4	3	0.0087	0.0155	-0.0069	0.0001	0.0003
X4	5	0.0096	0.0155	-0.0059	0.0001	0.0001
X40	intercept	-0.0625	-0.0973	0.0348	0.0061	0.0060
X40	1	1.7004	1.6617	0.0386	0.1224	0.1395
X40	3	1.1345	1.2061	-0.0716	0.0315	0.0375
X40	4	2.0046	2.0071	-0.0026	0.0251	0.0254
X41	intercept	0.6745	0.6442	0.0304	0.0064	0.0063
X41	1	0.8558	0.8965	-0.0407	0.0305	0.0396
X41	4	1.1188	1.1737	-0.0549	0.0136	0.0131
X41	5	0.1080	0.1115	-0.0035	0.0067	0.0083
X41	8	1.5853	1.5312	0.0542	0.0343	0.0860
X42	intercept	-0.3465	-0.3861	0.0396	0.0073	0.0063
X42	1	1.0643	1.1112	-0.0469	0.0765	0.0823
X42	3	1.3584	1.3578	0.0006	0.0563	0.0414
X42	4	1.9265	1.9322	-0.0057	0.0348	0.0390
X43	intercept	-0.5506	-0.5676	0.0171	0.0063	0.0047
X43	1	1.5271	1.4684	0.0586	0.0772	0.0760
X43	4	1.0304	0.9880	0.0424	0.0205	0.0262

Table B12 (continued)

X43	7	0.8139	0.9083	-0.0944	0.0187	0.0292
X44	intercept	1.2739	1.2731	0.0008	0.0070	0.0055
X44	1	2.4791	2.4444	0.0347	0.0815	0.0918
X44	3	0.6843	0.5770	0.1073	0.0293	0.0268
X44	4	1.0291	1.1443	-0.1152	0.0168	0.0261
X44	5	0.9690	0.9624	0.0067	0.0176	0.0175
X45	intercept	1.3015	1.2817	0.0197	0.0087	0.0081
X45	1	1.9142	1.9142	-0.0001	0.0489	0.0407
X45	3	0.6778	0.5760	0.1018	0.0245	0.0208
X45	4	1.3931	1.4458	-0.0526	0.0175	0.0166
X45	5	0.6013	0.6512	-0.0499	0.0133	0.0246
X46	intercept	0.2107	0.1810	0.0297	0.0086	0.0061
X46	1	1.4771	1.5012	-0.0242	0.1014	0.0837
X46	3	0.7512	0.6615	0.0897	0.0463	0.0246
X46	4	1.3117	1.4221	-0.1104	0.0235	0.0300
X46	5	0.7301	0.6914	0.0387	0.0240	0.0229
X47	intercept	1.1218	1.1475	-0.0257	0.0086	0.0040
X47	1	0.8989	0.8230	0.0759	0.0169	0.0185
X47	3	0.8239	0.7262	0.0977	0.0126	0.0162
X47	4	0.8677	0.9309	-0.0632	0.0216	0.0146
X47	5	0.4089	0.4980	-0.0891	0.0145	0.0182
X48	intercept	1.3036	1.2677	0.0359	0.0072	0.0074
X48	1	0.9507	1.0402	-0.0895	0.0182	0.0205
X48	2	0.1261	0.2199	-0.0937	0.0093	0.0098
X48	4	0.8206	0.7824	0.0382	0.0131	0.0208
X48	8	1.3177	1.1456	0.1722	0.0159	0.0220
X49	intercept	-0.1631	-0.1627	-0.0004	0.0069	0.0066
X49	1	1.6573	1.7614	-0.1041	0.1192	0.1057
X49	2	1.3587	1.3859	-0.0272	0.0242	0.0247
X49	4	1.1482	1.0950	0.0532	0.0222	0.0317
X5	intercept	0.3147	0.3522	-0.0374	0.0071	0.0055
X5	2	1.1134	1.1838	-0.0704	0.0321	0.0274
X5	3	0.9737	1.0430	-0.0693	0.0280	0.0385
X5	4	1.4105	1.3681	0.0424	0.0164	0.0254
X5	7	0.2198	0.1954	0.0243	0.0209	0.0173
X50	intercept	2.8884	2.8843	0.0041	0.0156	0.0195
X50	1	1.8900	1.8756	0.0144	0.0154	0.0190
X50	2	0.2780	0.1464	0.1316	0.0290	0.0110
X50	3	0.6155	0.5724	0.0432	0.0148	0.0218
X50	4	0.6923	0.9543	-0.2620	0.0311	0.0504

Table B12 (continued)

X50	5	0.7960	0.7013	0.0947	0.0225	0.0279
X51	intercept	-1.1277	-1.1247	-0.0031	0.0089	0.0115
X51	1	8.3366	7.1879	1.1487	1.0919	4.0637
X51	2	1.6895	1.7635	-0.0740	0.1195	0.1171
X51	3	5.6282	3.1346	2.4936	4.7873	0.9978
X51	4	1.5388	1.4199	0.1190	0.0678	0.1061
X52	intercept	1.8230	1.7407	0.0824	0.0107	0.0089
X52	1	1.8232	1.8676	-0.0443	0.0302	0.0325
X52	2	0.5946	0.5987	-0.0041	0.0192	0.0257
X52	4	1.6748	1.5827	0.0921	0.0184	0.0155
X52	7	0.7067	0.7366	-0.0299	0.0223	0.0175
X53	intercept	1.7513	1.8011	-0.0498	0.0062	0.0154
X53	1	1.5100	1.4617	0.0482	0.0229	0.0254
X53	2	0.4046	0.5335	-0.1289	0.0279	0.0377
X53	4	0.9163	0.9005	0.0158	0.0246	0.0189
X53	7	0.9835	0.9586	0.0249	0.0185	0.0138
X54	intercept	1.4748	1.4613	0.0135	0.0049	0.0049
X54	1	2.0702	2.0629	0.0073	0.0144	0.0128
X54	3	0.1411	0.1190	0.0221	0.0097	0.0075
X54	4	0.0961	0.0871	0.0090	0.0049	0.0042
X55	intercept	1.2439	1.2231	0.0208	0.0054	0.0063
X55	1	2.6864	2.5960	0.0905	0.0234	0.0230
X55	3	0.4401	0.4956	-0.0555	0.0182	0.0183
X55	4	0.3339	0.2599	0.0740	0.0225	0.0108
X56	intercept	0.6445	0.5938	0.0507	0.0038	0.0043
X56	1	2.8927	2.9286	-0.0359	0.0380	0.0384
X56	7	0.5388	0.4268	0.1120	0.0074	0.0080
X57	intercept	1.3187	1.3313	-0.0126	0.0086	0.0057
X57	1	2.5977	2.5362	0.0615	0.0279	0.0227
X57	4	0.1020	0.0859	0.0161	0.0073	0.0047
X57	5	0.1867	0.2395	-0.0528	0.0121	0.0165
X57	8	0.4348	0.4032	0.0316	0.0257	0.0231
X58	intercept	-1.3528	-1.3698	0.0170	0.0107	0.0106
X58	2	0.6233	0.7025	-0.0791	0.0655	0.0512
X58	3	0.3634	0.4011	-0.0377	0.0476	0.0558
X58	4	1.8550	1.7633	0.0917	0.0630	0.0576
X58	6	0.4240	0.4646	-0.0407	0.0939	0.0815
X59	intercept	-1.0957	-1.1123	0.0166	0.0100	0.0093
X59	2	0.5381	0.5574	-0.0193	0.0440	0.0302
X59	3	0.5812	0.6019	-0.0206	0.0410	0.0430

Table B12 (continued)

X59	4	0.9589	0.9368	0.0221	0.0422	0.0511
X59	6	0.3545	0.3195	0.0350	0.0445	0.0440
X6	intercept	1.1188	1.1430	-0.0242	0.0092	0.0081
X6	2	0.7898	0.8891	-0.0993	0.0199	0.0155
X6	3	0.3492	0.3246	0.0246	0.0265	0.0173
X6	4	0.9287	0.9365	-0.0079	0.0219	0.0181
X6	6	1.8020	1.7193	0.0827	0.0473	0.0594
X60	intercept	-0.7295	-0.7008	-0.0287	0.0084	0.0087
X60	2	1.0885	1.1308	-0.0423	0.0402	0.0409
X60	3	0.4031	0.4245	-0.0214	0.0377	0.0514
X60	4	1.1634	1.1389	0.0246	0.0351	0.0364
X60	6	0.1625	0.2017	-0.0392	0.0201	0.0313
X61	intercept	0.2050	0.1814	0.0236	0.0051	0.0078
X61	2	0.4053	0.4265	-0.0212	0.0202	0.0215
X61	3	0.3782	0.3441	0.0341	0.0200	0.0194
X61	4	0.8499	0.8551	-0.0053	0.0140	0.0146
X61	6	1.1615	1.1797	-0.0182	0.0328	0.0406
X62	intercept	-0.6533	-0.6269	-0.0264	0.0060	0.0066
X62	2	0.2149	0.2398	-0.0249	0.0180	0.0175
X62	3	0.9724	0.9646	0.0078	0.0329	0.0375
X62	4	0.7281	0.7112	0.0169	0.0146	0.0219
X62	6	0.2111	0.2884	-0.0773	0.0238	0.0343
X63	intercept	2.0405	2.0524	-0.0119	0.0129	0.0136
X63	2	0.0986	0.0836	0.0150	0.0055	0.0047
X63	3	0.1224	0.1255	-0.0031	0.0077	0.0096
X63	4	0.6627	0.7216	-0.0589	0.0158	0.0236
X63	5	0.0659	0.0692	-0.0033	0.0031	0.0034
X63	6	1.9832	2.0222	-0.0389	0.0142	0.0160
X64	intercept	0.0556	0.0851	-0.0295	0.0072	0.0057
X64	2	0.9629	0.9581	0.0048	0.0369	0.0344
X64	3	0.9234	0.9823	-0.0589	0.0324	0.0534
X64	4	1.5740	1.6611	-0.0871	0.0271	0.0350
X64	6	0.6602	0.5391	0.1211	0.1002	0.0687
X65	intercept	-1.6607	-1.5772	-0.0836	0.0104	0.0106
X65	2	1.2700	1.3395	-0.0695	0.0401	0.0550
X65	3	0.5564	0.5565	-0.0002	0.0467	0.0557
X66	intercept	-0.1500	-0.0793	-0.0707	0.0054	0.0056
X66	2	1.7064	1.7207	-0.0143	0.0105	0.0125
X67	intercept	0.9690	0.9618	0.0072	0.0055	0.0076
X67	1	0.6310	0.5408	0.0902	0.0179	0.0200

Table B12 (continued)

X67	2	0.1964	0.1954	0.0010	0.0112	0.0136
X67	3	0.0645	0.0557	0.0088	0.0026	0.0020
X67	4	0.8534	0.8789	-0.0254	0.0115	0.0170
X67	6	1.2127	1.3375	-0.1248	0.0203	0.0282
X67	10	0.0671	0.0789	-0.0118	0.0029	0.0045
X68	intercept	0.4109	0.4372	-0.0263	0.0049	0.0064
X68	1	1.4676	1.5131	-0.0455	0.0257	0.0238
X68	2	1.4079	1.3558	0.0522	0.0115	0.0103
X69	intercept	0.4647	0.5383	-0.0736	0.0061	0.0042
X69	1	1.9760	1.9535	0.0225	0.0499	0.0530
X69	2	1.6920	1.6955	-0.0035	0.0132	0.0097
X7	intercept	-1.9264	-1.9051	-0.0213	0.0108	0.0098
X7	2	1.9862	1.8641	0.1221	0.0562	0.0391
X70	intercept	0.6829	0.7291	-0.0462	0.0062	0.0083
X70	1	1.3166	1.3189	-0.0023	0.0221	0.0265
X70	2	1.6757	1.6616	0.0141	0.0109	0.0130
X71	intercept	0.6173	0.6475	-0.0302	0.0057	0.0044
X71	1	2.1729	2.1607	0.0121	0.0613	0.0404
X71	2	1.0959	1.0559	0.0400	0.0170	0.0242
X71	10	0.7201	0.7483	-0.0282	0.0216	0.0310
X72	intercept	0.7026	0.7302	-0.0276	0.0065	0.0065
X72	1	1.7300	1.8047	-0.0747	0.0299	0.0314
X72	2	1.2047	1.1494	0.0553	0.0253	0.0181
X72	4	0.4361	0.4745	-0.0385	0.0174	0.0198
X73	intercept	-0.3829	-0.3706	-0.0122	0.0050	0.0064
X73	1	1.4983	1.4822	0.0161	0.0414	0.0357
X73	2	0.8054	0.6463	0.1591	0.0237	0.0334
X73	10	0.2007	0.3483	-0.1476	0.0191	0.0297
X74	intercept	0.0975	0.0994	-0.0019	0.0050	0.0059
X74	1	0.7611	0.7745	-0.0135	0.0195	0.0182
X74	2	0.1761	0.2730	-0.0968	0.0115	0.0185
X74	4	0.3160	0.3049	0.0111	0.0100	0.0154
X74	10	0.7299	0.6334	0.0966	0.0160	0.0195
X75	intercept	-1.2985	-1.2921	-0.0064	0.0066	0.0053
X75	1	2.0192	2.5595	-0.5404	0.4516	0.6472
X75	3	0.3233	0.3476	-0.0244	0.0377	0.0358
X75	5	2.4811	2.2825	0.1986	0.0871	0.0703
X76	intercept	-0.0340	-0.1014	0.0674	0.0055	0.0070
X76	1	1.1499	1.1447	0.0052	0.0457	0.0389
X76	4	0.8675	0.8331	0.0344	0.0165	0.0157

Table B12 (continued)

X76	10	0.9739	0.9548	0.0191	0.0169	0.0174
X77	intercept	-0.9512	-0.9321	-0.0191	0.0060	0.0051
X77	1	1.4446	1.6158	-0.1712	0.1197	0.1596
X77	5	0.5893	0.7154	-0.1262	0.0966	0.0491
X77	10	1.5445	1.3639	0.1806	0.1120	0.0736
X78	intercept	0.8544	0.8804	-0.0259	0.0079	0.0060
X78	1	0.9787	1.1226	-0.1439	0.0254	0.0381
X78	3	0.1721	0.2595	-0.0873	0.0133	0.0186
X78	4	0.7783	0.8050	-0.0267	0.0160	0.0137
X78	7	0.2289	0.2556	-0.0267	0.0190	0.0222
X78	9	6.1630	1.4230	4.7399	2.5538	0.4920
X78	10	0.4183	0.3969	0.0214	0.0325	0.0191
X79	intercept	0.7324	0.7573	-0.0249	0.0048	0.0065
X79	1	1.7599	1.7506	0.0093	0.0396	0.0259
X79	3	0.3417	0.3013	0.0404	0.0193	0.0202
X79	5	1.5338	1.5640	-0.0303	0.0191	0.0149
X8	intercept	-0.8411	-0.7501	-0.0910	0.0051	0.0084
X8	2	2.5850	2.6107	-0.0257	0.0303	0.0285
X80	intercept	-0.7265	-0.7128	-0.0137	0.0038	0.0038
X80	1	1.7993	1.9023	-0.1030	0.1048	0.1310
X80	5	1.7632	1.7508	0.0124	0.0158	0.0230
X81	intercept	0.2350	0.2476	-0.0126	0.0035	0.0048
X81	1	1.6579	1.6582	-0.0003	0.0359	0.0388
X81	5	1.6322	1.6267	0.0055	0.0090	0.0117
X82	intercept	0.1592	0.1750	-0.0158	0.0044	0.0042
X82	1	1.3732	1.3488	0.0244	0.0271	0.0381
X82	5	0.6714	0.7527	-0.0813	0.0367	0.0411
X82	10	0.8910	0.8346	0.0564	0.0336	0.0331
X83	intercept	-0.8780	-0.8821	0.0041	0.0031	0.0025
X83	1	2.5000	2.8813	-0.3813	0.2583	0.2557
X83	6	9.1361	7.8100	1.3262	0.4544	1.9597
X84	intercept	1.0713	1.1035	-0.0322	0.0040	0.0053
X84	1	0.7470	0.7412	0.0058	0.0149	0.0190
X84	3	0.3641	0.4327	-0.0686	0.0156	0.0127
X84	4	0.4363	0.5186	-0.0823	0.0101	0.0124
X84	6	0.7503	0.7743	-0.0240	0.0915	0.0546
X84	7	0.0952	0.0645	0.0307	0.0063	0.0034
X84	8	0.3784	0.2855	0.0929	0.0839	0.0386
X85	intercept	0.0541	0.0413	0.0128	0.0069	0.0056
X85	1	1.0103	0.8951	0.1152	0.0376	0.0468

Table B12 (continued)

X85	3	0.7264	0.6673	0.0591	0.0242	0.0190
X85	4	0.6221	0.6206	0.0016	0.0138	0.0110
X85	8	0.4746	0.5925	-0.1179	0.0481	0.0446
X85	10	0.1755	0.1913	-0.0157	0.0146	0.0134
X86	intercept	1.4161	1.4027	0.0134	0.0037	0.0034
X86	1	0.7519	0.7472	0.0046	0.0175	0.0128
X86	3	0.0914	0.0772	0.0142	0.0042	0.0040
X86	4	0.0378	0.0403	-0.0026	0.0011	0.0013
X86	6	0.5899	0.5887	0.0011	0.0212	0.0122
X9	intercept	1.3552	1.4222	-0.0670	0.0079	0.0106
X9	2	0.5425	0.6355	-0.0930	0.0291	0.0293
X9	4	0.9217	1.0323	-0.1106	0.0203	0.0244
X9	6	2.4810	2.3674	0.1136	0.0330	0.0374
X9	10	0.1768	0.1626	0.0142	0.0171	0.0125

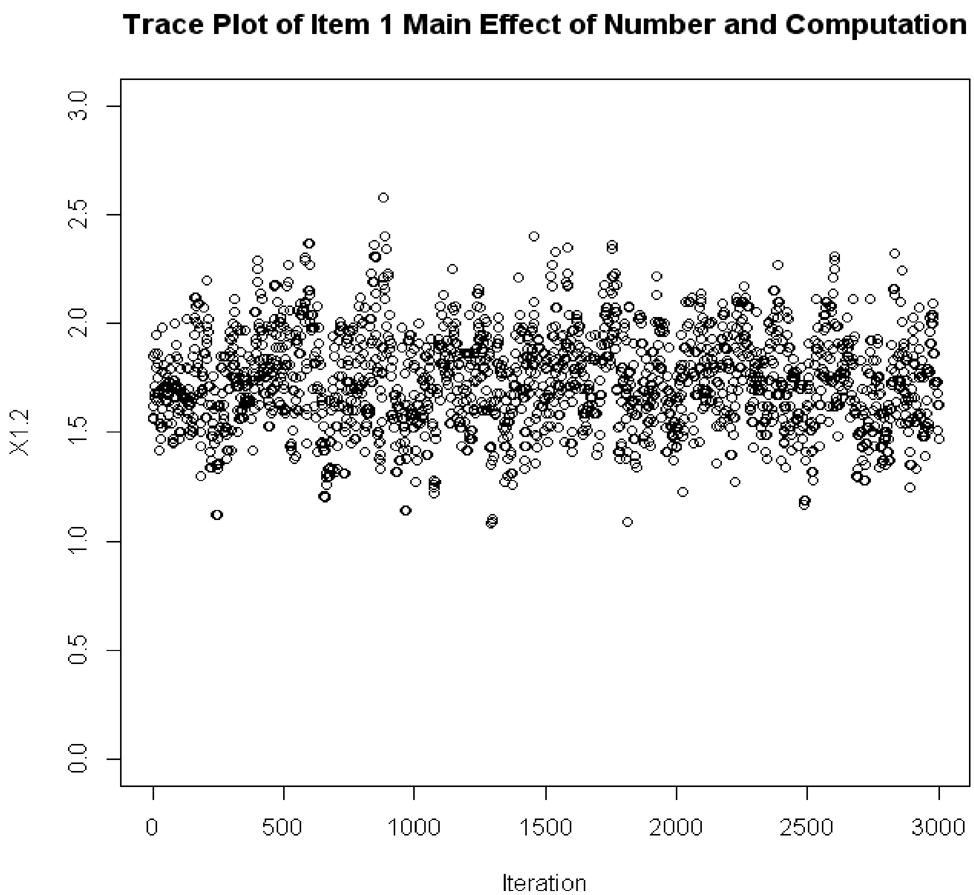

Figure B 1. The trace plot for CRUM main effect of the Number and Computation attribute for item 1 is stable, with no apparent pattern in the post-burnin chain.

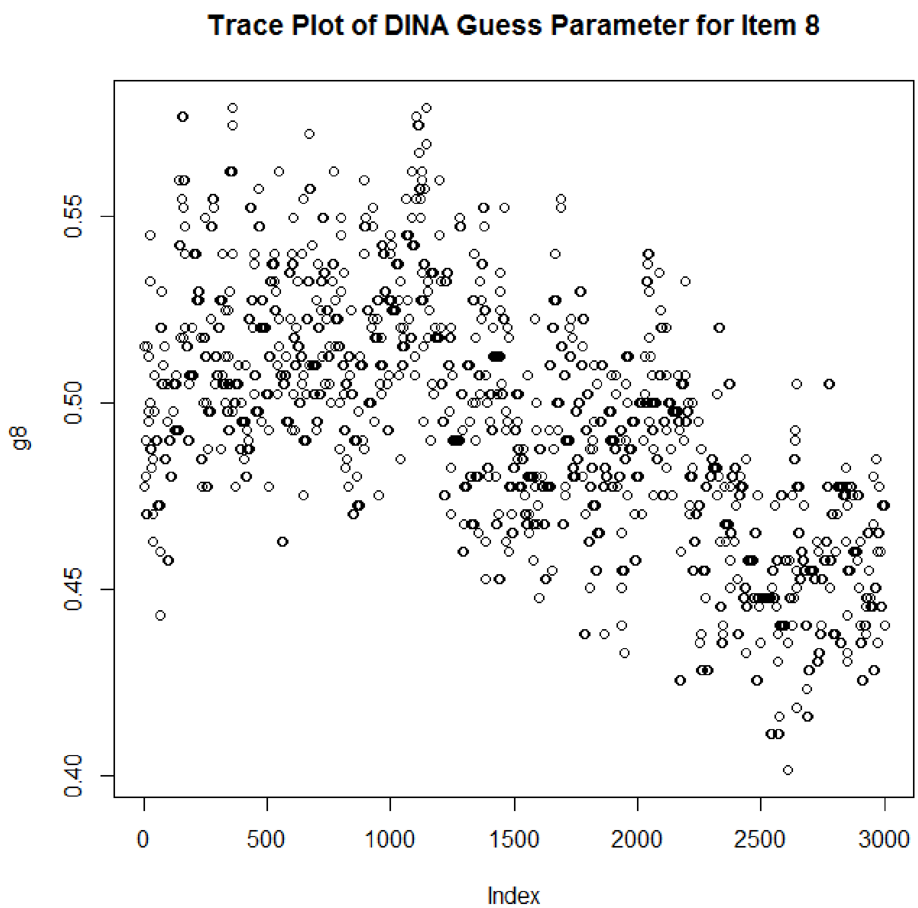

Figure B 2. The trace plot for DINA guess parameter for item 8 is unstable, with a decreasing trend for later steps in the chain

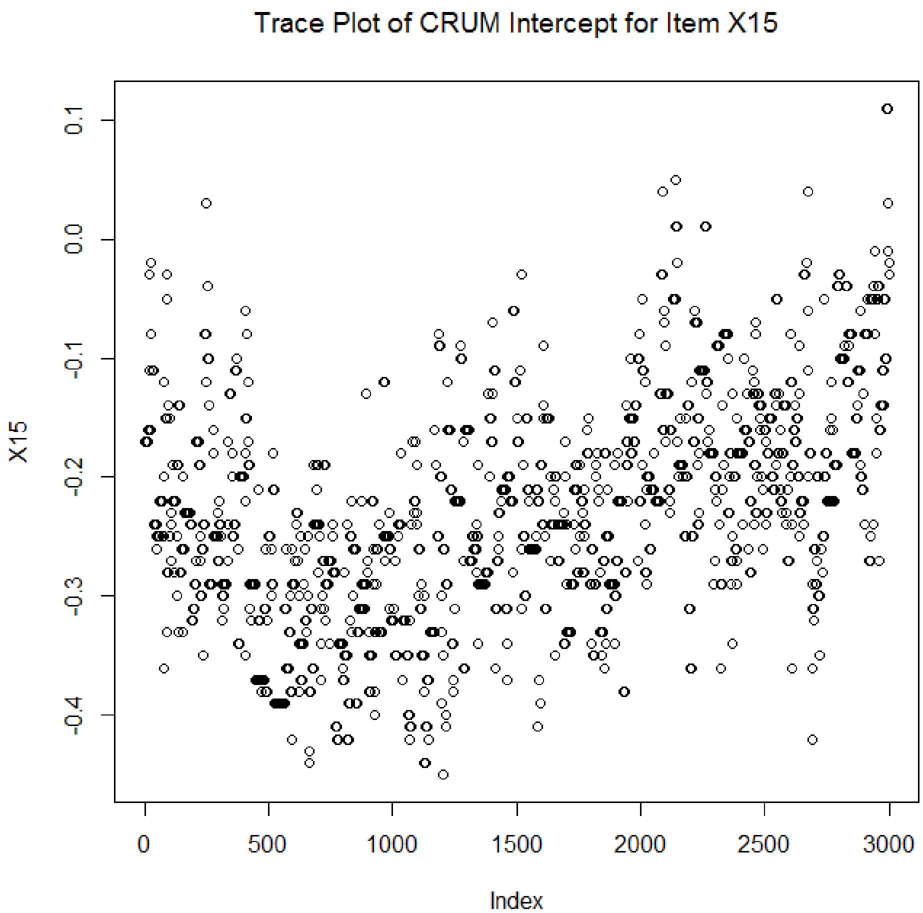

Figure B 3. The trace plot for CRUM intercept for item 15 is unstable, with an increasing trend for later steps in the chain

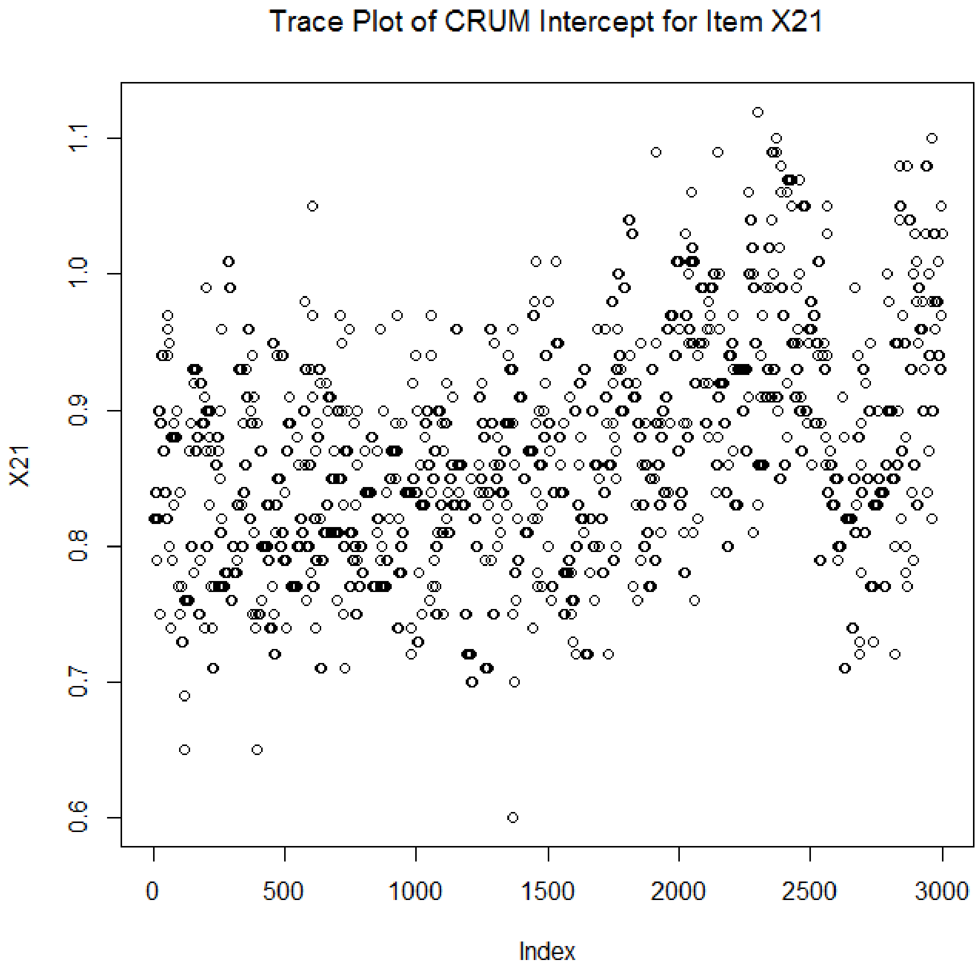

Figure B 4. The trace plot for CRUM intercept for item 21 is unstable, with an increasing trend for later steps in the chain

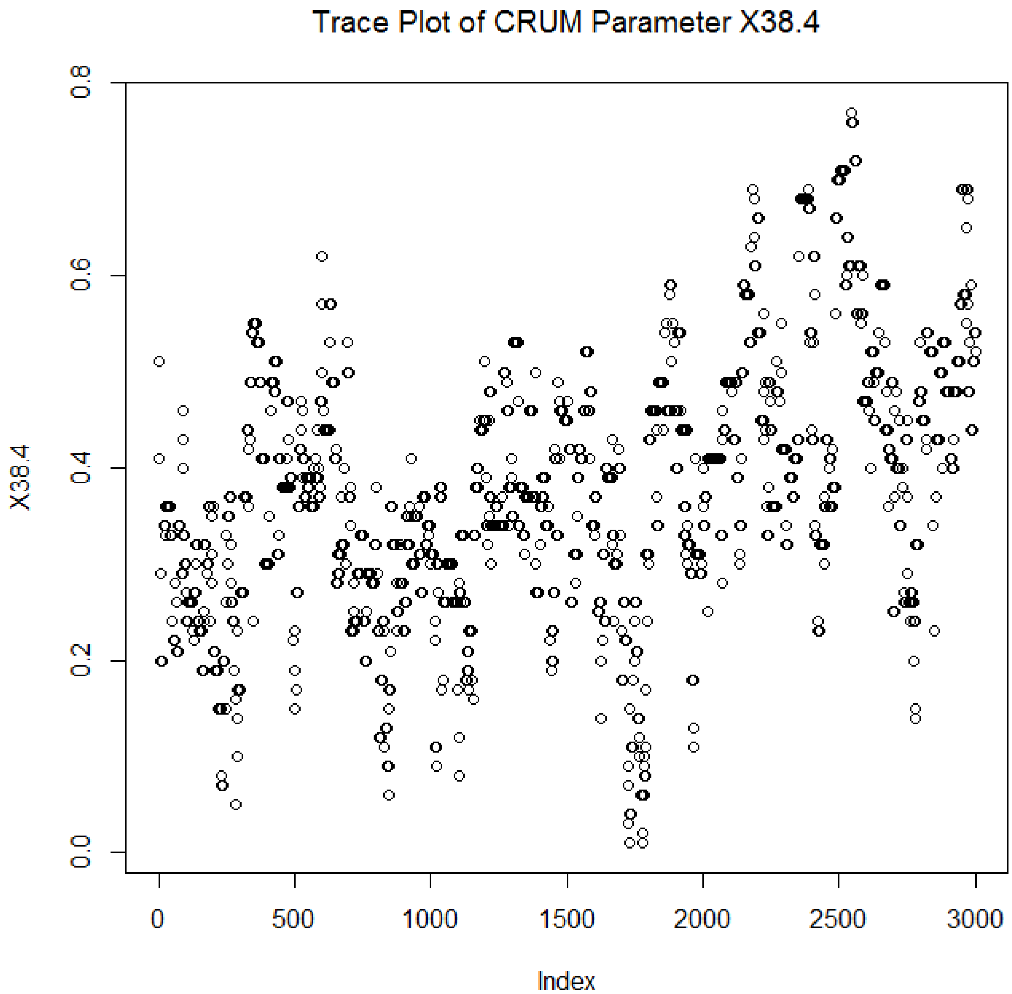

Figure B 5. The trace plot for CRUM main effect parameter for item 15 attribute 4 is unstable, with an increasing trend for later steps in the chain

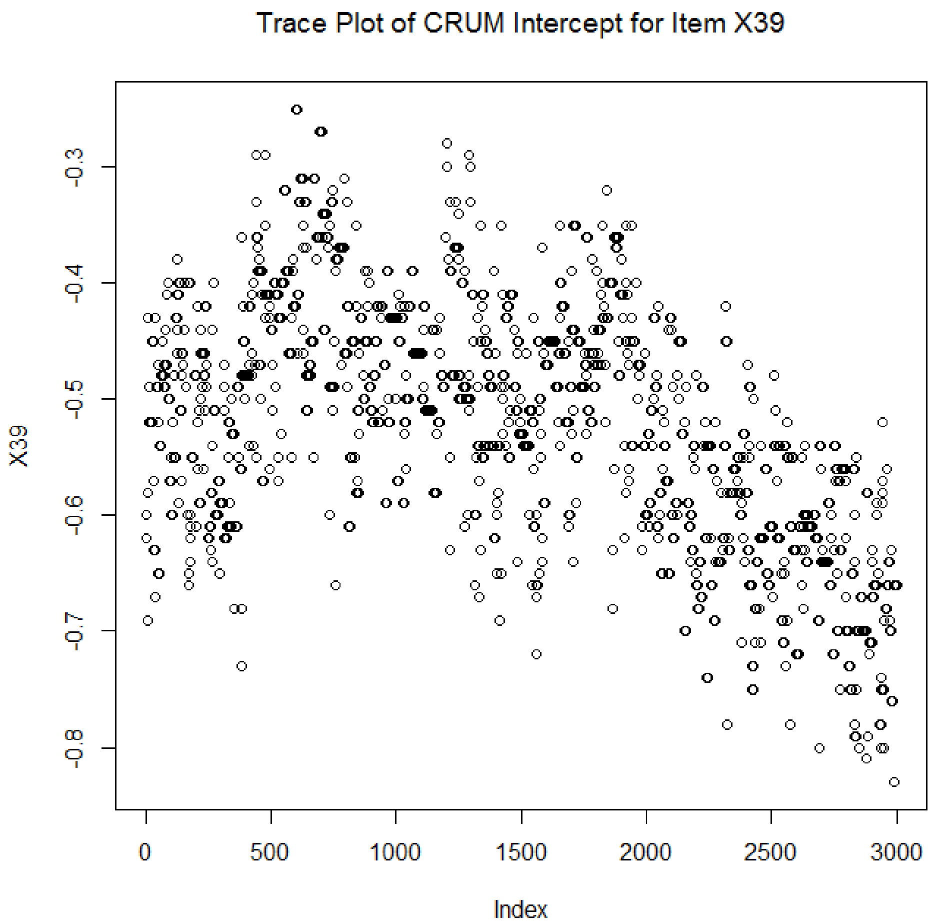

Figure B 6. The trace plot for CRUM intercept for item 39 is unstable, with a decreasing trend for later steps in the chain

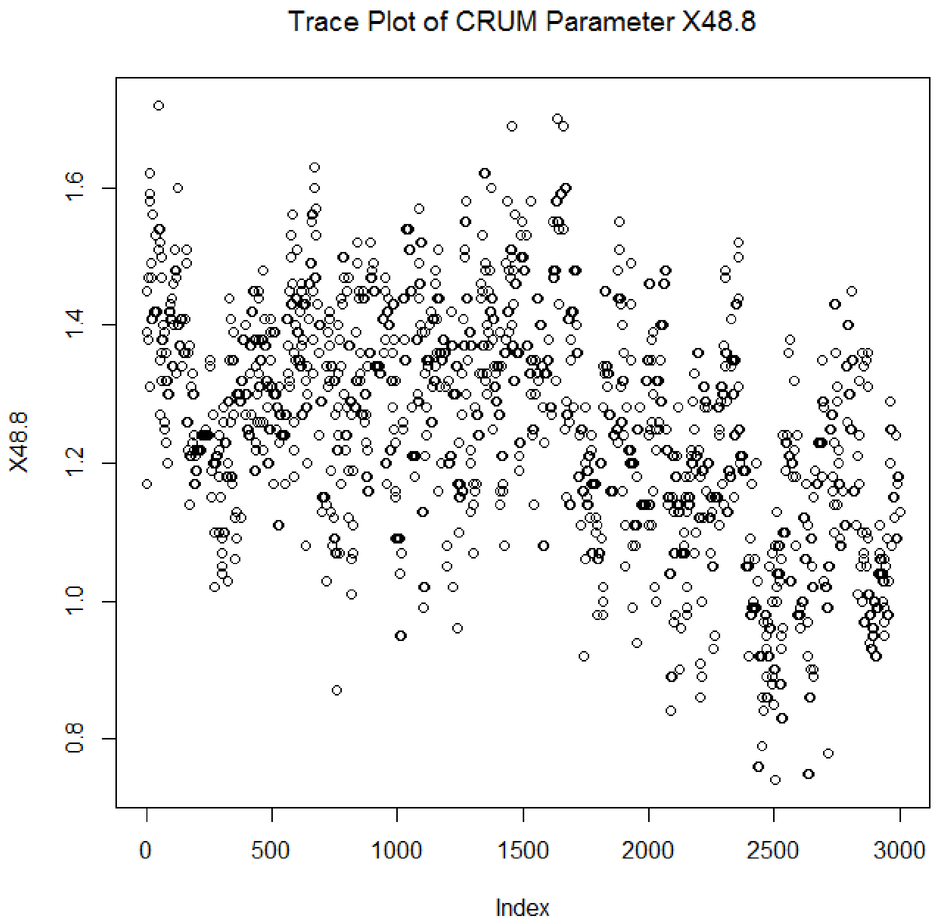

Figure B 7. The trace plot for CRUM main effect parameter for item 48 attribute 8 is unstable, with a decreasing trend for later steps in the chain

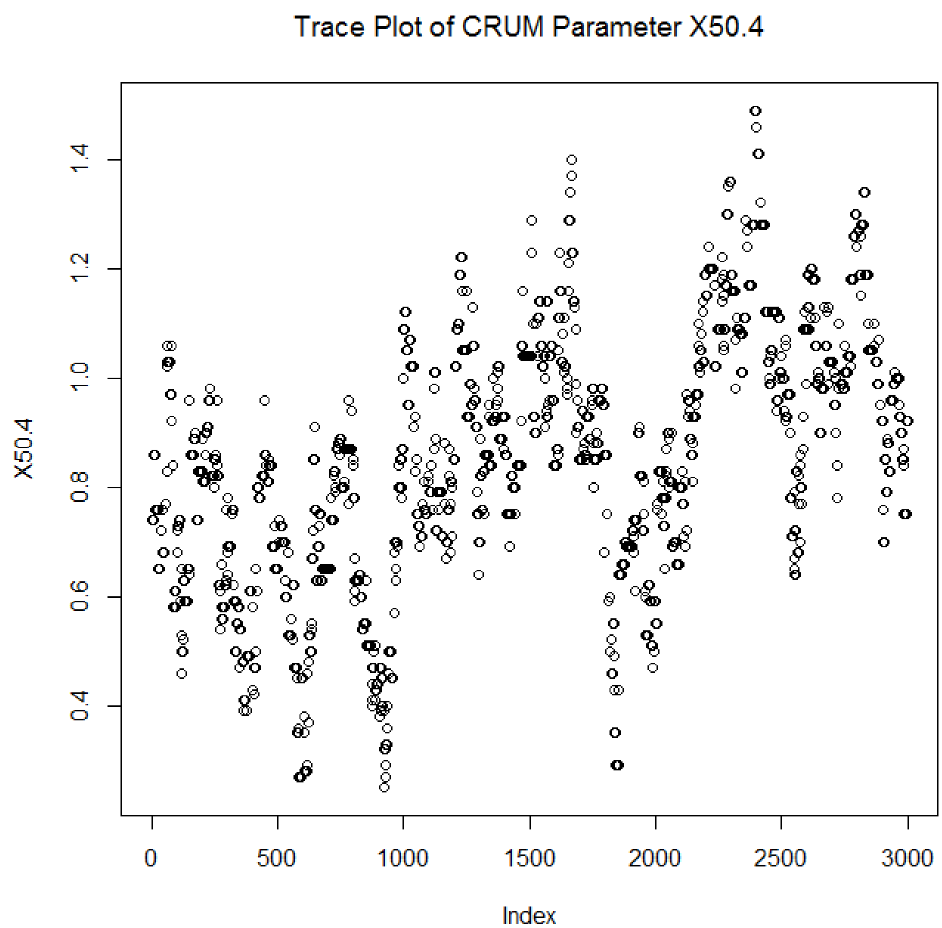

Figure B 8. The trace plot for CRUM main effect parameter for item 50 attribute 4 is unstable, with periodicity throughout the chain

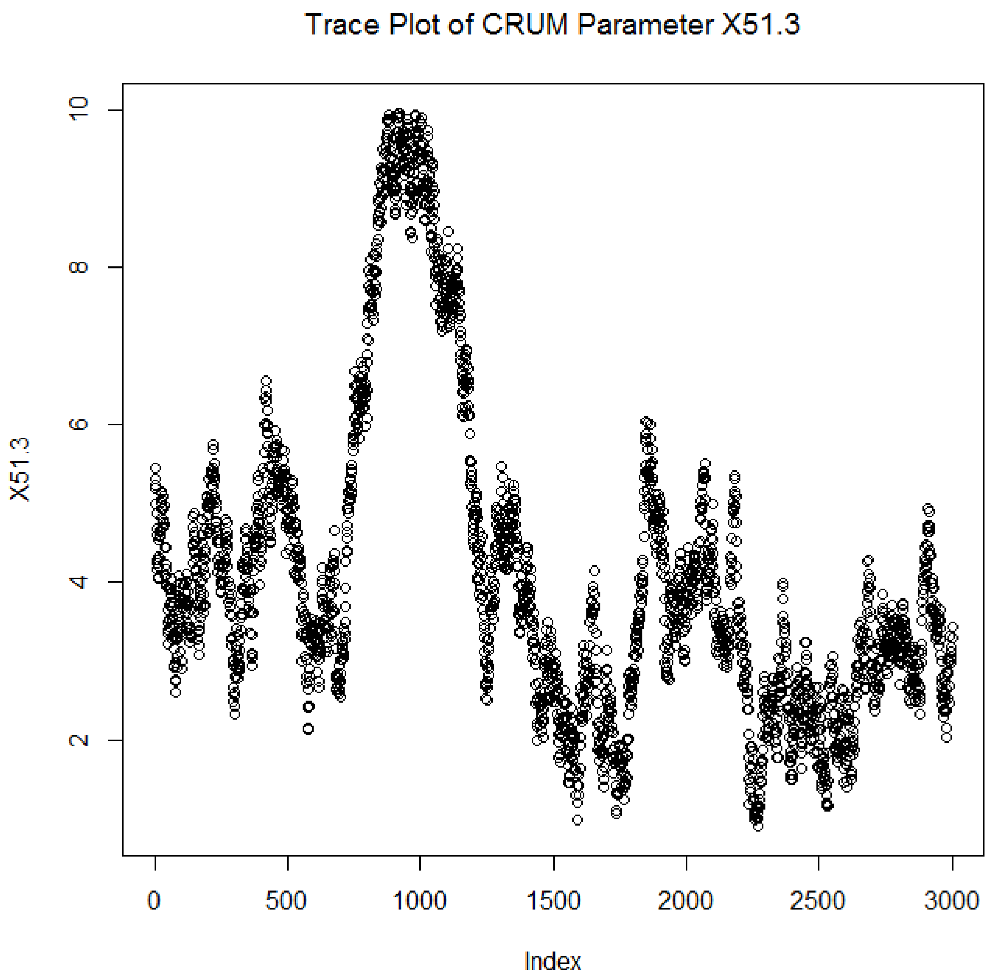

Figure B 9. The trace plot for CRUM main effect parameter for item 51 attribute 3 is unstable, with periodicity throughout the chain

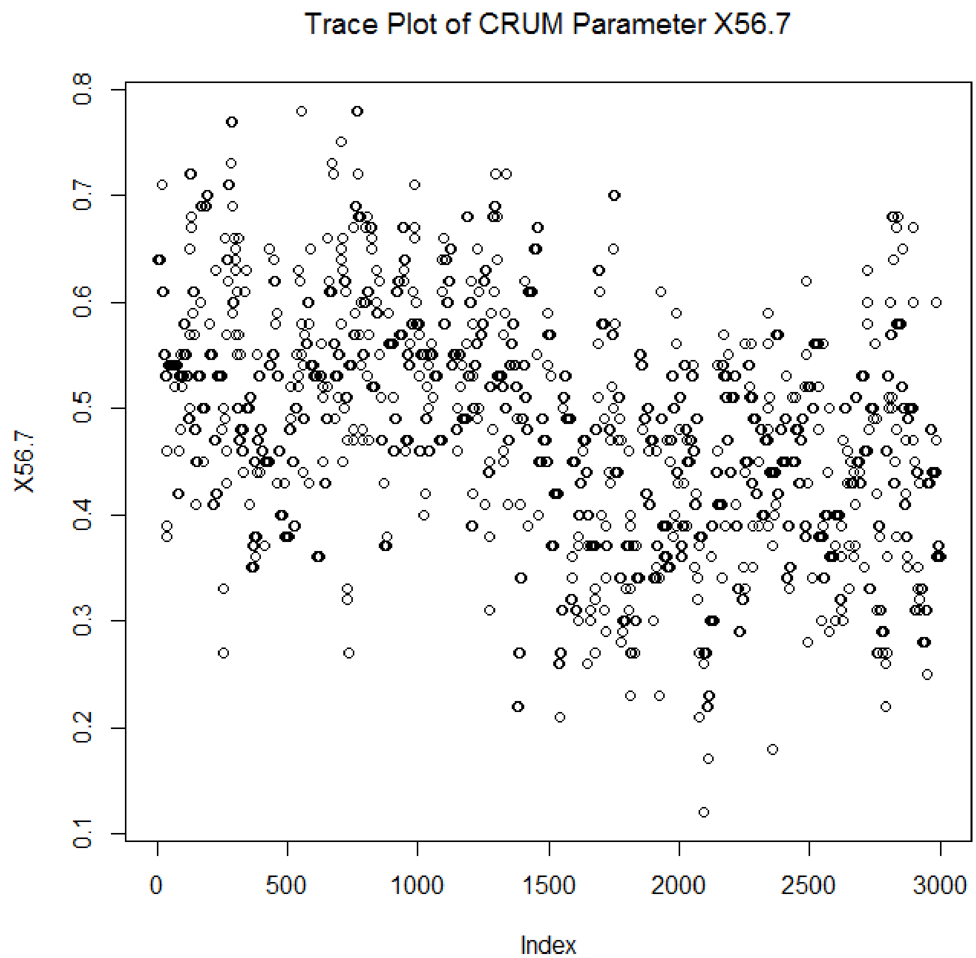

Figure B 10. The trace plot for CRUM main effect parameter for item 56 attribute 7 is unstable, with a decreasing trend for later steps in the chain

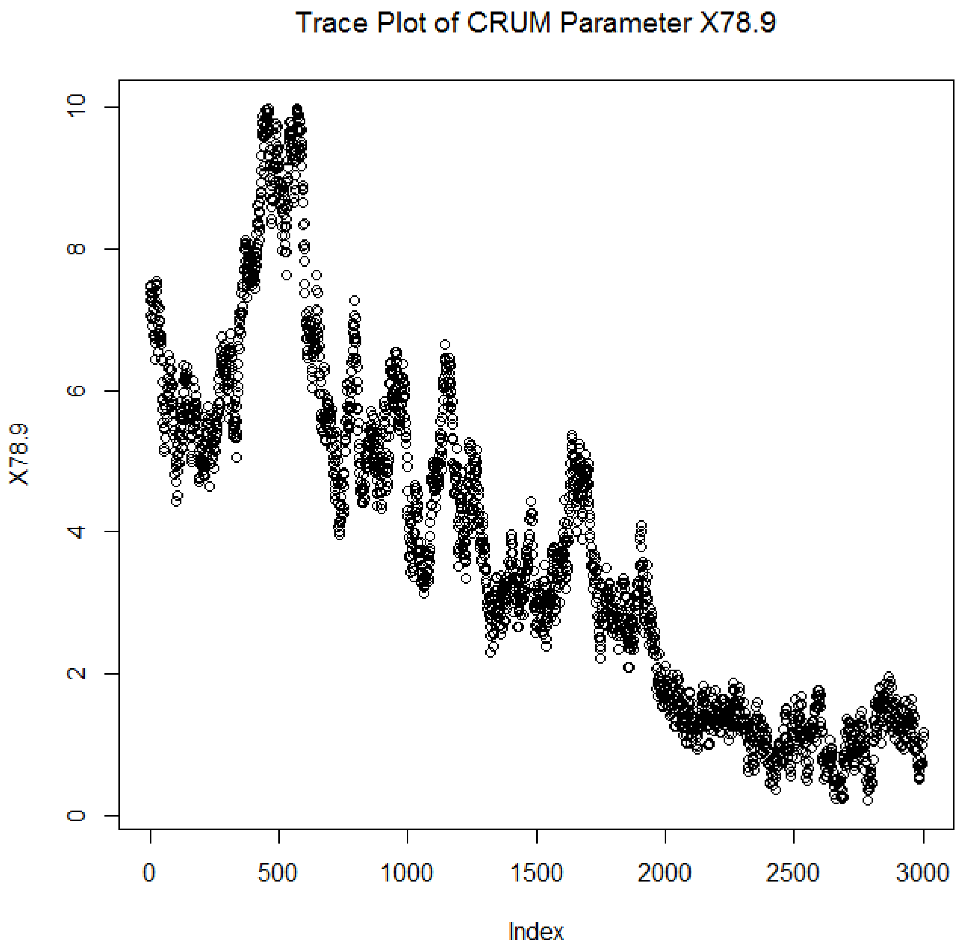

Figure B 11. The trace plot for CRUM main effect parameter for item 78 attribute 9 is unstable, with periodicity around a decreasing trend over the chain

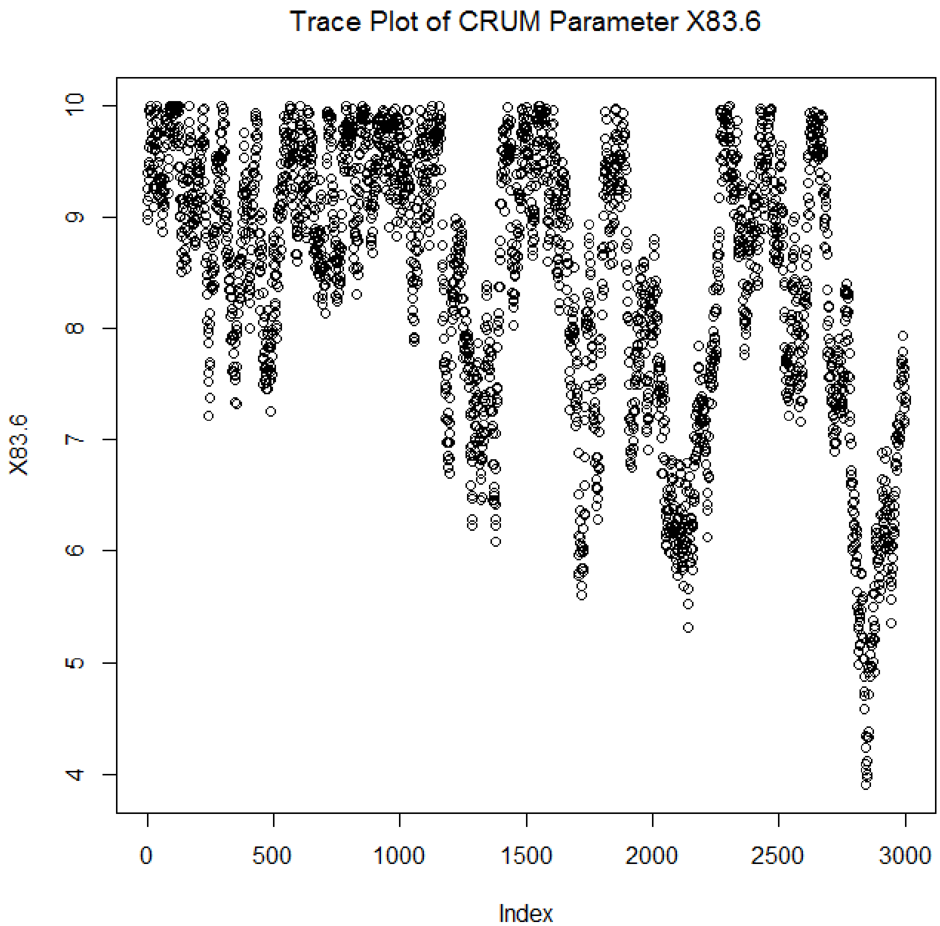

Figure B 12. The trace plot for CRUM main effect parameter for item 83 attribute 6 is unstable, with periodicity over the chain

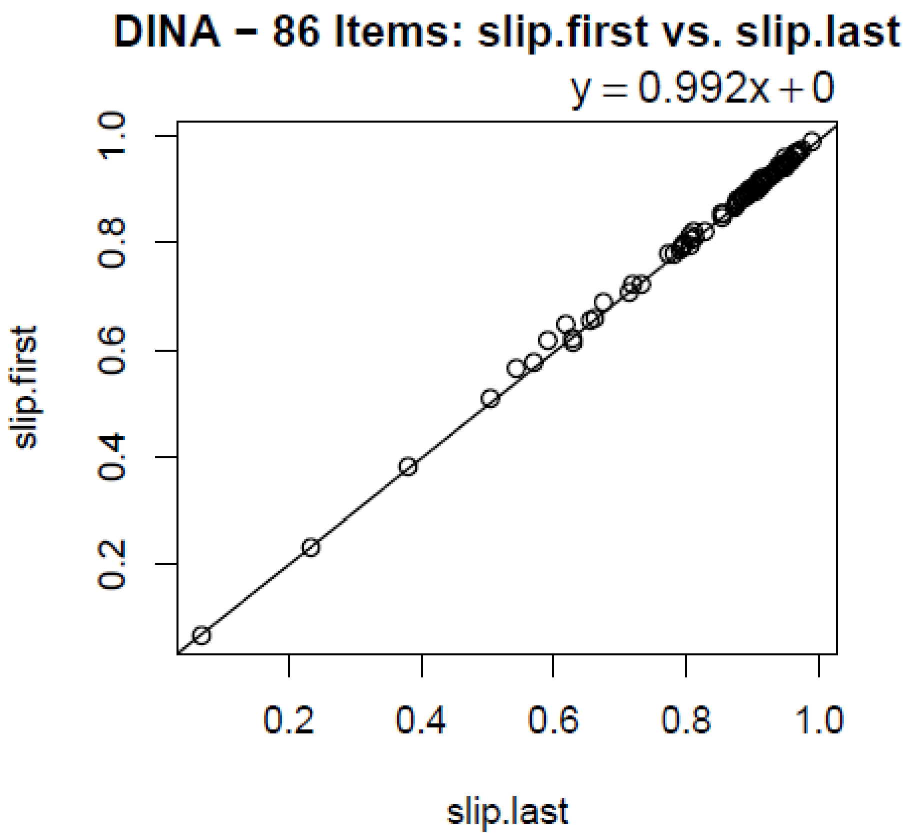

Figure B 13. A scatterplot of the DINA slip parameter means from the first and last Geweke windows (p1 = p2 = 0.40). The fitted linear equation indicates overall convergence of the parameters.

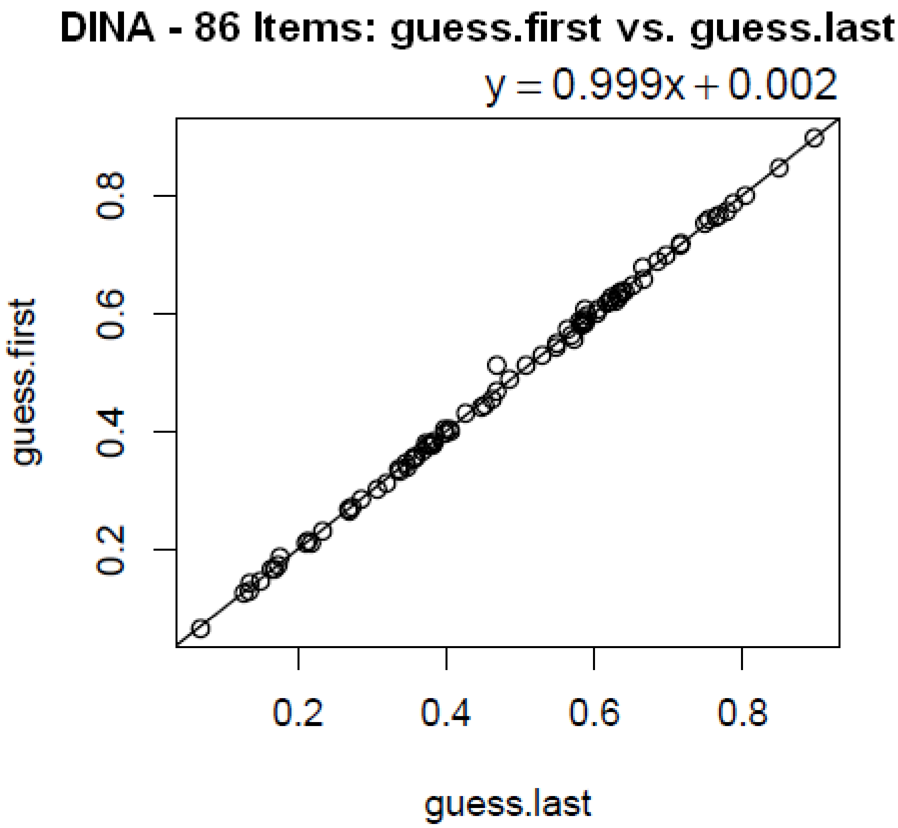

Figure B 14. A scatterplot of the DINA guess parameter means from the first and last Geweke windows (p1 = p2 = 0.40). The fitted linear equation indicates overall convergence of the parameters.

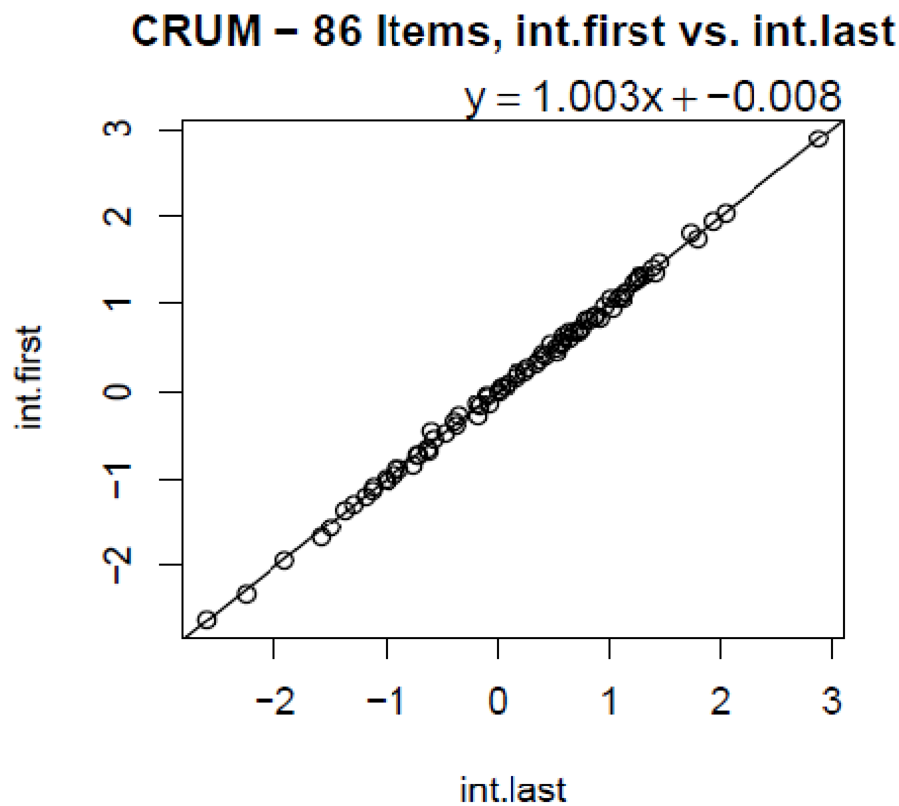

Figure B 15. A scatterplot of the CRUM intercept means from the first and last Geweke windows (p1 = p2 = 0.40). The fitted linear equation indicates overall convergence of the parameters.

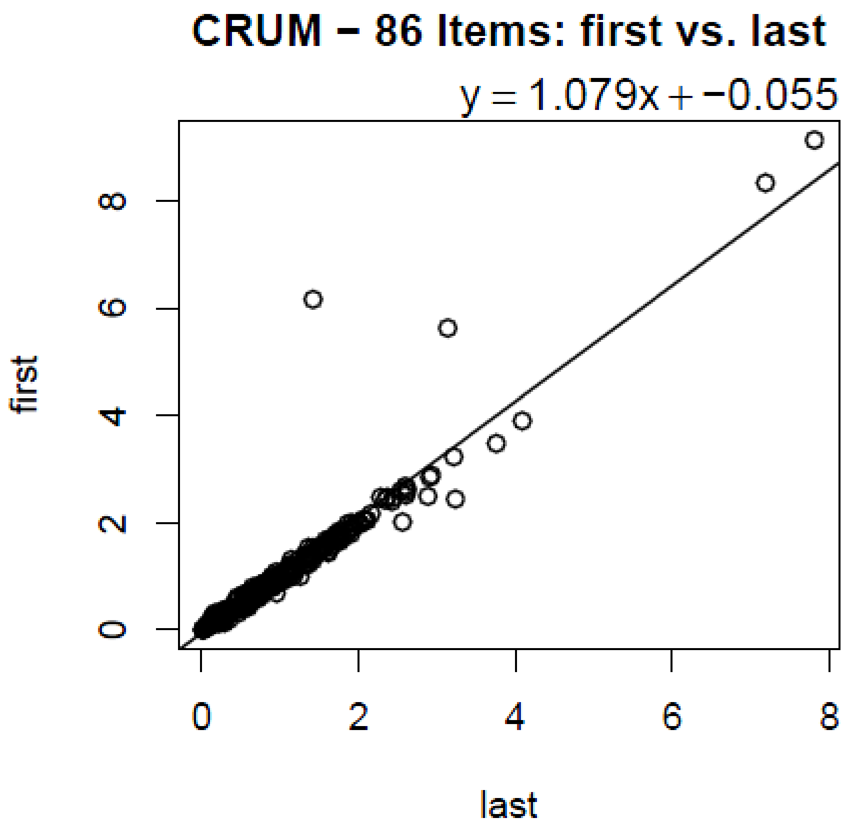

Figure B 16. scatterplot of the CRUM main effects parameter means from the first and last Geweke windows (p1 = p2 = 0.40). The fitted linear equation indicates overall convergence of the parameters.

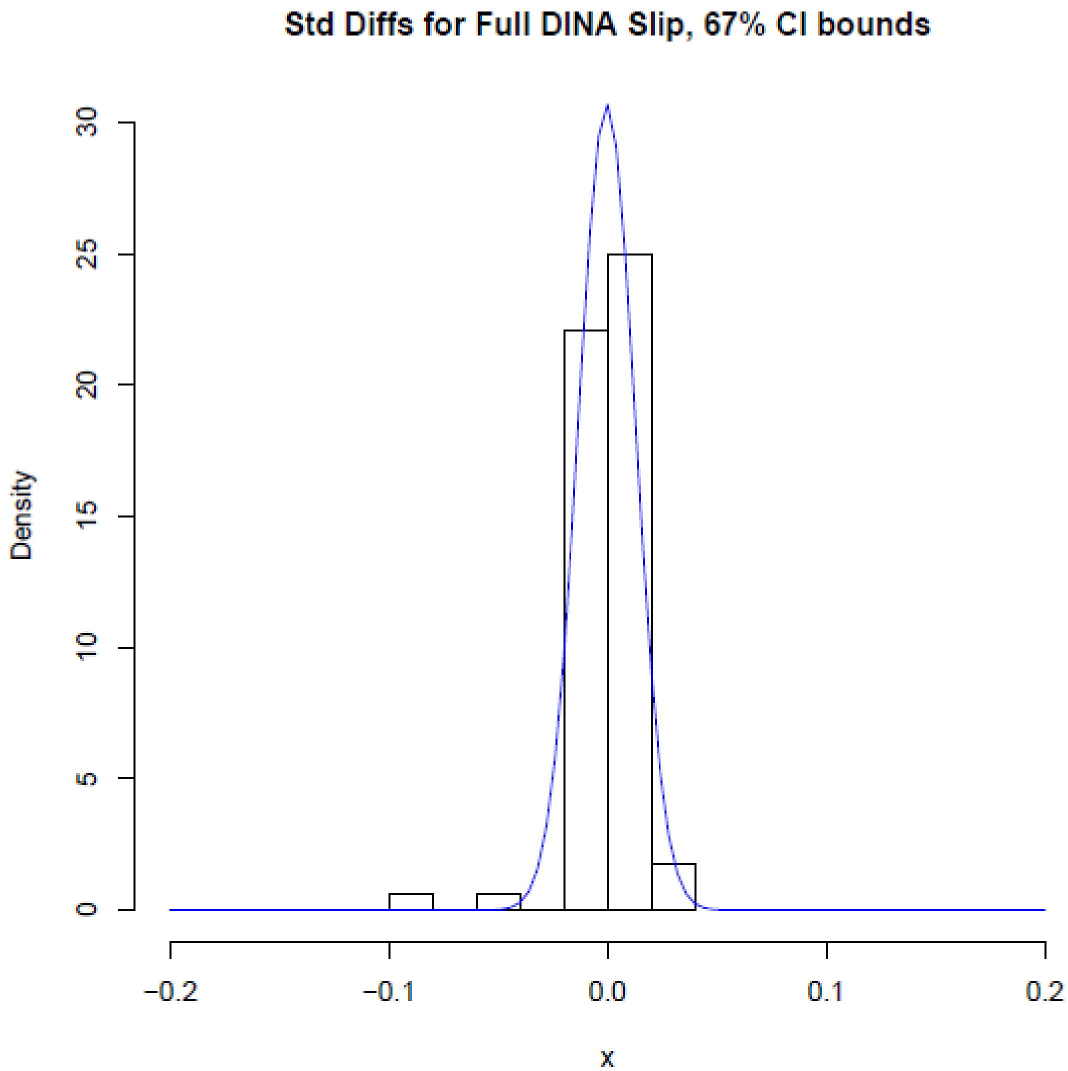

Figure B 17. Distribution of standardized differences for the slip parameters of the DINA model. The overlaid curve is the fitted normal distribution of the difference quantiles. The CI bounds are not visible because they lie at values beyond the scale of the x-axis

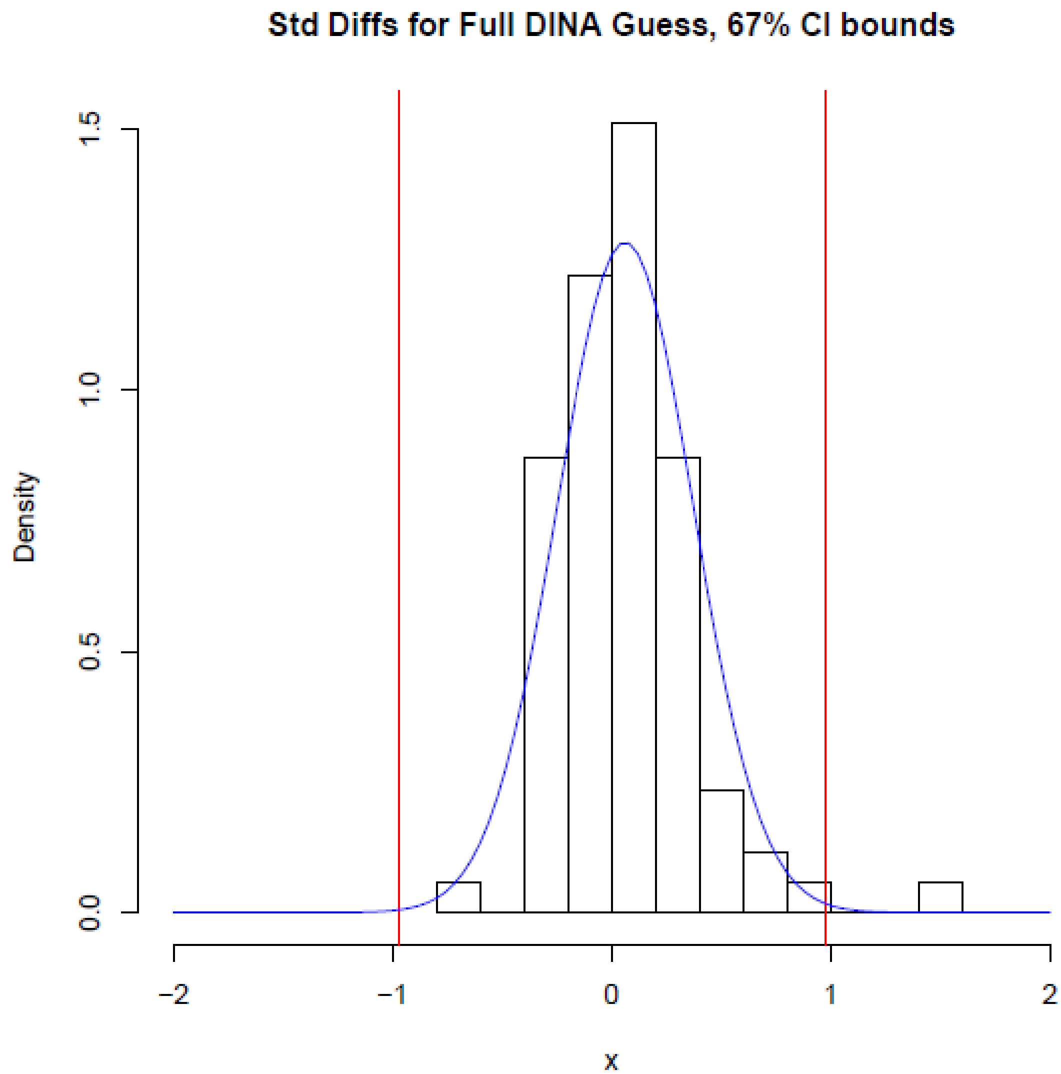

Figure B 18. Distribution of standardized differences for the guess parameters of the DINA model. Vertical lines represent 67% confidence limits, the overlaid curve is the fitted normal distribution of the difference quantiles.

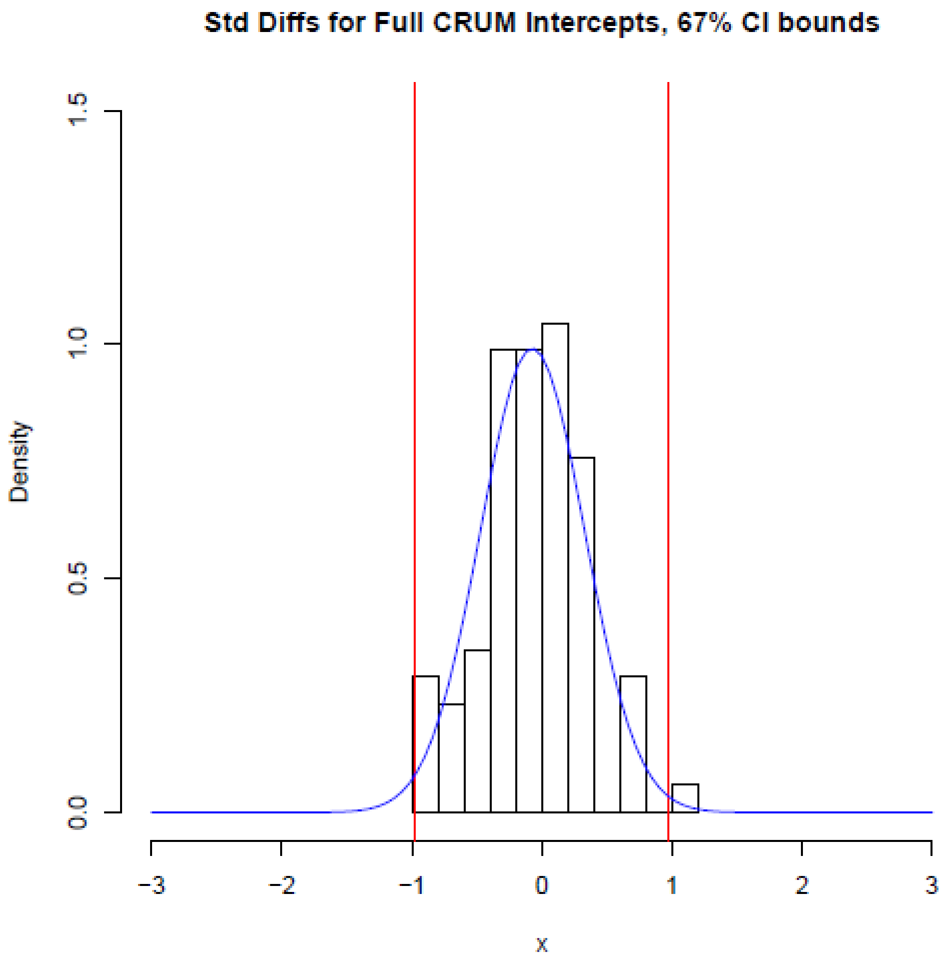

Figure B 19. Distribution of standardized differences for the intercept parameters of the CRUM model. Vertical lines represent 67% confidence limits, the overlaid curve is the fitted normal distribution of the difference quantiles

.

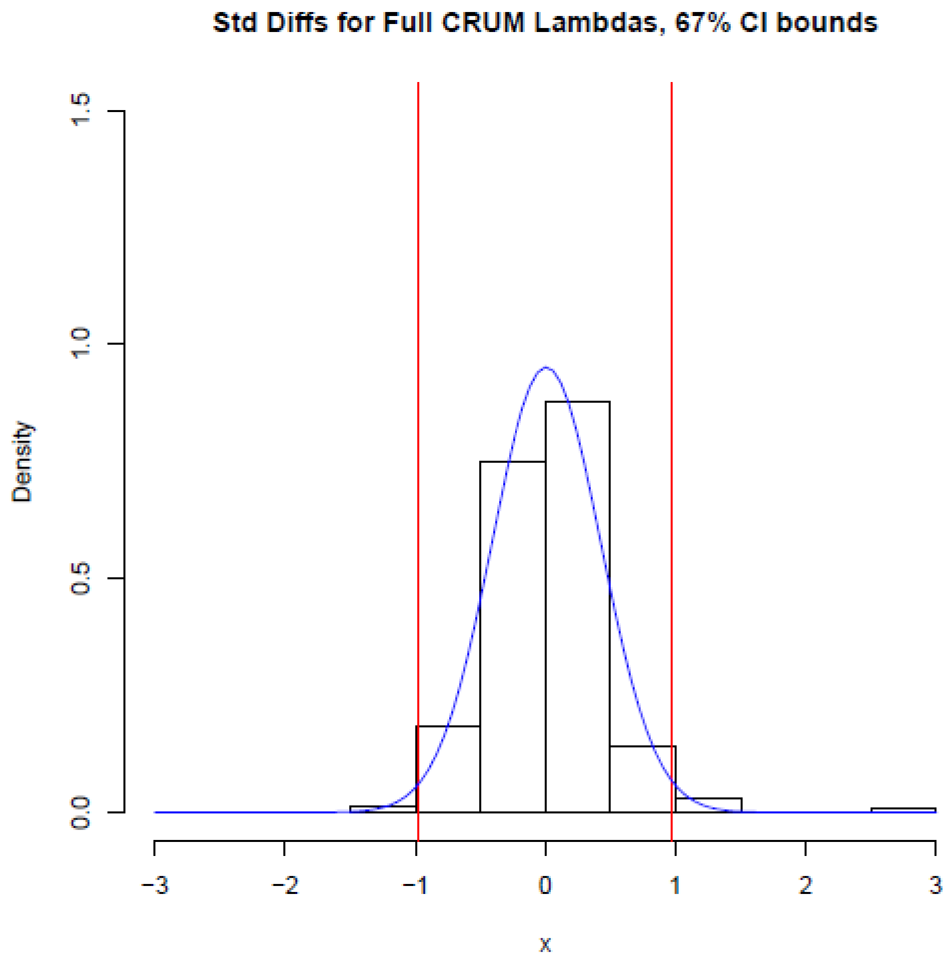

Figure B 20. Distribution of standardized differences for the main effects parameters of the CRUM model. Vertical lines represent 67% confidence limits, the overlaid curve is the fitted normal distribution of the difference quantiles.

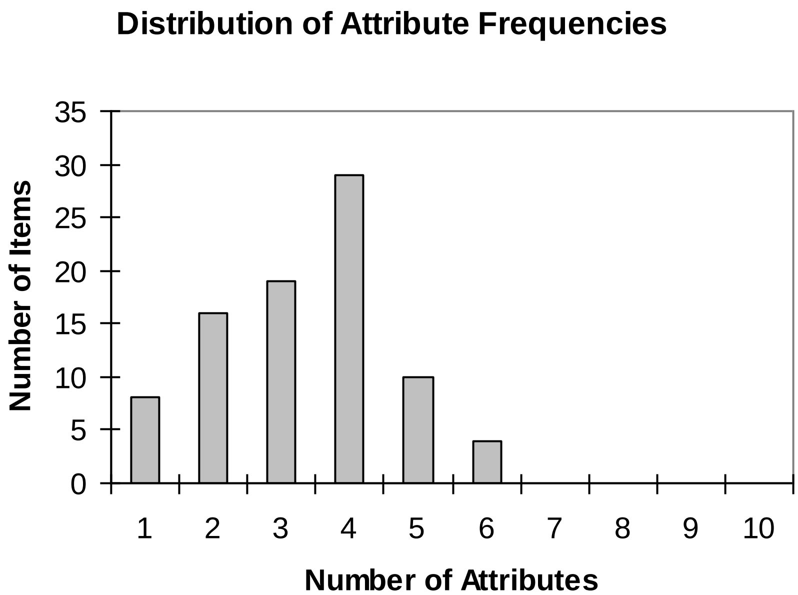

Figure B 21. Distribution of the attribute frequencies among items. No items have more than six attributes

APPENDIX C

REDUCED PARAMETER SPACE RESULTS AND CONVERGENCE DIAGNOSTICS

Table C 1. Summary of reduced item set's facility and biserial correlation from CTT analysis

Statistic	Proportion Answered Correctly	Biserial Correlation
Count	61	61
Minimum	33.00	0.39
Maximum	87.80	0.74
Mean	67.47	0.58
Standard Error	1.80	0.01
Median	69.20	0.57
Mode	56.40	0.67
Standard Deviation	14.04	0.08
Sample Variance	197.16	0.01
Range	54.80	0.36

*Table C 2. **Q**-matrix for reduced item and attribute set*

Item	Test Position	Number and Computation	Contextual Encoding	Encode Diagram	Generate Equations or Plausible Values
5	0	1	1	0	0
6	0	1	1	0	1
9	0	1	1	0	1
11	1	1	0	0	0
13	1	1	1	0	0
15	1	1	0	0	0
17	1	1	1	0	0
18	1	1	1	0	0
19	1	0	0	1	1
20	1	0	0	1	0
21	1	0	0	0	0
22	1	0	0	0	0
23	1	0	1	1	0
24	1	0	1	0	1
25	1	0	1	0	1
26	1	0	1	1	0
27	1	0	1	0	1
28	1	0	1	0	1
29	1	0	1	0	0
30	1	0	1	0	1
31	0	0	1	0	0
32	0	0	0	0	0
33	0	0	1	0	1
35	0	0	1	0	0
39	0	0	1	0	0
40	1	0	1	0	0
41	1	0	1	1	0
42	1	0	1	0	0
43	1	0	1	0	0
44	1	0	1	1	0
45	1	0	1	1	0
46	1	0	1	1	0
47	1	0	1	1	0
48	1	1	1	0	0
49	1	1	1	0	0
50	1	1	1	1	0
52	1	1	1	0	0
53	1	1	1	0	0
54	1	0	1	0	0
55	1	0	1	0	0
56	1	0	0	0	0
57	1	0	1	1	0
61	0	1	1	0	1
63	0	1	1	1	1

Table C2 (continued)

64	0	1	1	0	1
67	1	1	1	0	1
68	1	1	0	0	0
69	1	1	0	0	0
70	1	1	0	0	0
71	1	1	0	0	0
72	1	1	1	0	0
73	1	1	0	0	0
76	1	0	1	0	0
78	1	0	1	0	0
79	1	0	0	1	0
80	1	0	0	1	0
81	1	0	0	1	0
82	1	0	0	1	0
83	1	0	0	0	1
84	1	0	1	0	1
85	1	0	1	0	0

Table C2 (continued)

Item	Visualization	Number of Subgoals	Bottom-up Processing
5	1	0	0
6	0	0	0
9	0	0	1
11	0	0	0
13	0	1	0
15	0	0	0
17	0	0	0
18	0	0	0
19	0	0	1
20	0	0	0
21	1	0	0
22	0	0	1
23	0	0	0
24	0	0	1
25	0	0	0
26	1	0	1
27	0	1	0
28	0	1	0
29	0	0	1
30	0	1	0
31	0	0	0
32	0	0	0
33	1	0	0
35	0	0	0
39	0	0	0
40	0	0	0
41	0	1	0
42	0	0	0
43	1	0	0
44	0	0	0
45	0	0	0
46	0	0	0
47	0	0	0
48	0	1	0
49	0	0	0
50	0	0	0
52	1	0	0
53	1	0	0
54	0	0	0
55	0	0	0
56	1	0	0
57	0	1	0
61	0	0	0
63	0	0	0
64	0	0	0

Table C2 (continued)

67	0	0	1
68	0	0	0
69	0	0	0
70	0	0	0
71	0	0	1
72	0	0	0
73	0	0	1
76	0	0	1
78	1	0	1
79	0	0	0
80	0	0	0
81	0	0	0
82	0	0	1
83	0	0	0
84	1	1	0
85	0	1	1

Table C 3. ANOVA results for DINA slip parameters

Source	df	SS	MS	F
Model	1	0.71696	0.71696	64,061 *
Error	59	0.00066	0.00001	--
Total	60	0.71762	--	--

* $p < 0.01$

Coefficient	Estimate	Std. Error	t-value
(Intercept)	-0.002090	0.00355	-0.587
s2.Red	1.001577	0.003957	253.103 *

* $p < 0.01$

Table C 4. ANOVA results for DINA guess parameters

Source	df	SS	MS	F
Model	1	1.58817	1.58817	67,270*
Error	59	0.00139	0.00002	--
Total	60	1.58956	--	--

* $p < 0.01$

Coefficient	Estimate	Std. Error	t-value
(Intercept)	-0.001329	0.002174	-0.611
g2.Red	1.000577	0.003858	259.364 *

* $p < 0.01$

Table C 5. ANOVA results for CRUM intercepts

Source	df	SS	MS	F
Model	1	39.195	39.195	64,933 *
Error	59	0.036	0.001	--
Total	60	39.231	--	--

* $p < 0.01$

Coefficient	Estimate	Std. Error	*t*-value
(Intercept)	0.001125	0.003609	0.312
int.Red	1.000044	0.003925	254.820 *

* $p < 0.01$

Table C 6. ANOVA results for CRUM main effects

Source	df	SS	MS	F
Model	1	142.996	142.996	13,096 *
Error	59	1.933	0.011	--
Total	60	144.929	--	--

* $p < 0.01$

Coefficient	Estimate	Std. Error	*t*-value
(Intercept)	-0.005179	0.012900	-0.402
me.Red	1.008551	0.008813	114.438 *

* $p < 0.01$

Table C 7. Standardized differences for DINA parameters

Item	Slip	Guess
X5	-0.0501	-0.0764
X6	-0.1341	-0.0925
X9	-0.1324	-0.0128
X11	0.1621	0.0644
X13	0.1035	0.0420
X15	0.0302	-0.0290
X17	-0.0004	-0.0600
X18	0.0254	-0.0983
X19	-0.0501	0.1537
X20	-0.1200	-0.0769
X21	-0.1446	-0.5128
X22	-0.0658	0.1347
X23	-0.0589	0.0496
X24	-0.0513	0.1222
X25	-0.0111	-0.0754
X26	0.0880	0.0724
X27	0.0349	-0.0757
X28	-0.0357	-0.1085
X29	0.0167	0.0227
X30	-0.2670	-0.2997
X31	-0.2314	-0.2569
X32	0.0204	0.0204
X33	-0.2220	-0.2649
X35	0.5528	0.4554
X39	-0.7671	-0.7966
X40	0.2607	0.1587
X41	-0.1591	-0.1960
X42	0.1440	-0.0418
X43	-0.2683	-0.4586
X44	0.0938	0.0377
X45	-0.2285	-0.2122
X46	-0.1284	-0.0046
X47	-0.2205	-0.1234
X48	-0.3836	-0.3671
X49	-0.1136	-0.1844
X50	0.2104	0.1537
X52	-0.0590	-0.0829
X53	-0.5460	-0.5704

Table C7 (continued)

X54	0.0446	0.0143
X55	0.5156	0.4964
X56	-0.2760	-0.3394
X57	-0.0987	-0.0829
X61	0.2505	0.3118
X63	0.3630	0.3172
X64	0.1208	0.1420
X67	0.1677	0.2286
X68	-0.2124	-0.2168
X69	-0.0151	0.0236
X70	-0.1428	-0.1766
X71	-0.1042	0.0432
X72	-0.2320	-0.2433
X73	-0.0907	-0.1101
X76	-0.2080	-0.0143
X78	-0.0833	-0.0585
X79	-0.2407	-0.1426
X80	0.1073	0.1097
X81	0.2985	0.2153
X82	0.0879	0.1032
X83	0.3573	0.1568
X84	-0.0037	0.0198
X85	-0.3129	-0.3256

Table C 8. Standardized differences for CRUM intercepts

Item	Standardized Difference
X5	-0.1014
X6	0.0198
X9	0.5098
X11	-0.0147
X13	-0.1453
X15	-0.4408
X17	-0.1643
X18	-0.1399
X19	-0.0914
X20	-0.5222
X21	0.1512
X22	0.3109
X23	-0.4909
X24	0.1472
X25	0.4067
X26	0.0180
X27	0.4573
X28	0.0407
X29	-0.3057
X30	0.0037
X31	0.3353
X32	-0.0071
X33	0.3921
X35	0.4126
X39	0.5309
X40	0.3373
X41	0.0263
X42	0.3273
X43	-0.2048
X44	-0.3601
X45	0.0159
X46	0.1233
X47	-0.0912
X48	-0.0763
X49	-0.1602
X50	0.0161
X52	0.0252

Table C8 (continued)

X53	0.1312
X54	0.2983
X55	0.2449
X56	-0.0822
X57	0.2328
X61	-0.1686
X63	-0.1927
X64	-0.1153
X67	0.1217
X68	-0.4885
X69	-0.1405
X70	-0.5948
X71	-0.1254
X72	0.1412
X73	-0.1367
X76	0.4125
X78	0.1323
X79	-0.2263
X80	-0.1389
X81	-0.4003
X82	0.2152
X83	0.3770
X84	-0.1763
X85	-0.0689

Table C 9. Standardized differences for CRUM main effects

Parameter	Standardized Difference
X5.7	0.4038
X5.4	-0.2848
X5.2	-0.0711
X6.6	-0.7737
X6.4	-0.1189
X6.2	0.6324
X9.10	-0.2965
X9.6	-0.4306
X9.4	0.1048
X9.2	0.7062
X11.2	0.4817
X11.1	-0.1196
X13.8	-0.6448
X13.4	-0.2121
X13.2	0.2892
X13.1	0.4872
X15.2	0.1808
X15.1	-0.1195
X17.4	0.3049
X17.2	-0.3027
X17.1	-0.0597
X18.4	0.3358
X18.2	-0.0999
X18.1	-0.1800
X19.10	-0.5399
X19.6	0.1306
X19.5	0.1995
X19.1	0.5478
X20.5	-0.4842
X20.1	0.3004
X21.7	0.3652
X21.1	0.1895
X22.10	-1.0524
X22.1	0.8789
X23.5	0.3676
X23.4	-0.4167
X23.1	0.1544

Table C9 (continued)

X24.10	0.2748
X24.6	0.3743
X24.4	-0.5420
X24.1	0.0491
X25.6	0.4719
X25.4	-0.1771
X25.1	-0.0346
X26.10	0.4623
X26.7	-0.2949
X26.5	-0.0665
X26.4	-0.3955
X26.1	0.1240
X27.8	-0.1822
X27.6	0.5159
X27.4	-0.2159
X27.1	0.0868
X28.8	-1.0004
X28.6	0.5789
X28.4	-0.2947
X28.1	0.4235
X29.10	0.2040
X29.4	-0.8537
X29.1	0.2193
X30.8	-0.0754
X30.6	0.6193
X30.4	-0.8273
X30.1	-0.0367
X31.4	0.1537
X33.7	-0.2634
X33.6	0.0857
X33.4	0.2862
X35.4	0.1225
X39.4	0.1793
X40.4	0.0288
X40.1	0.0353
X41.8	0.6063
X41.5	-0.1370
X41.4	-0.3245
X41.1	-0.3290
X42.4	0.4413

Table C9 (continued)

X42.1	-0.3332
X43.7	-0.5021
X43.4	-0.0141
X43.1	0.3532
X44.5	-0.0821
X44.4	-0.1735
X44.1	-0.3764
X45.5	-0.1114
X45.4	0.0355
X45.1	-0.1678
X46.5	0.2589
X46.4	-0.0525
X46.1	-0.2005
X47.5	0.4094
X47.4	-0.4622
X47.1	0.0484
X48.8	-0.0802
X48.4	0.1324
X48.2	-0.4047
X48.1	0.2053
X49.4	0.0544
X49.2	0.0888
X49.1	-0.0697
X50.5	0.2491
X50.4	-0.4507
X50.2	0.4383
X50.1	-0.3971
X52.7	0.6095
X52.4	-0.0734
X52.2	-0.4263
X52.1	-0.0925
X53.7	-1.0411
X53.4	0.2911
X53.2	0.8472
X53.1	-0.1111
X54.4	-0.0028
X54.1	0.3142
X55.4	-0.1021
X55.1	0.5301
X56.7	-0.2259

Table C9 (continued)

X56.1	-0.0480
X57.8	-1.0134
X57.5	0.4049
X57.4	-0.0823
X57.1	0.6580
X61.6	0.4113
X61.4	-0.2712
X61.2	-0.1841
X63.6	-0.2846
X63.5	0.2419
X63.4	-0.8756
X63.2	0.8130
X64.6	0.6914
X64.4	-0.1858
X64.2	-0.3739
X67.10	0.1258
X67.6	-0.0510
X67.4	-0.5839
X67.2	0.5149
X67.1	0.1314
X68.2	-0.1886
X68.1	0.1857
X69.2	0.4837
X69.1	-0.2179
X70.2	-0.3224
X70.1	0.1237
X71.10	-0.1447
X71.2	0.3579
X71.1	-0.2717
X72.4	0.2982
X72.2	0.1715
X72.1	-0.3866
X73.10	0.4064
X73.2	-0.1220
X73.1	-0.4410
X76.10	0.2781
X76.4	0.1128
X76.1	-0.3980
X78.10	1.2970
X78.7	-1.2280

Table C9 (continued)

X78.4	-0.0264
X78.1	-0.2471
X79.5	-0.1469
X79.1	-0.0162
X80.5	0.1520
X80.1	-0.1146
X81.5	-0.3557
X81.1	0.0377
X82.10	0.0809
X82.5	0.1441
X82.1	-0.1348
X83.6	0.2727
X83.1	0.1005
X84.8	0.8848
X84.7	-0.0606
X84.6	-0.2844
X84.4	-0.6188
X84.1	-0.1128
X85.10	-0.6569
X85.8	1.4809
X85.4	-0.0955
X85.1	-0.9361

Table C 10. Probabilities of successful item completion for items flagged parameters

Flagged Parameter	Model	Pr[X=1] First	Pr[X=1] Last	Pr[X=1] Overall	Difference
X22.10	CRUM	0.9973	0.9982	0.9979	-0.0009
X28.8	CRUM	0.9853	0.9885	0.9872	-0.0032
X53.7	CRUM	0.9955	0.9962	0.9960	-0.0008
X57.8	CRUM	0.9893	0.9910	0.9903	-0.0017
X78.10	CRUM	0.9832	0.9783	0.9803	0.0048
X78.7	CRUM	0.9775	0.9819	0.9803	-0.0044
X85.8	CRUM	0.9559	0.9333	0.9425	0.0226

Table C 11. Flagged items and attributes and their properties

Item	Model	Flagged Attribute	Facility	Item's Attribute Saturation	Attribute's Test Representation
X22	CRUM	10	0.812	2	13
X28	CRUM	8	0.711	4	9
X53	CRUM	7	0.541	4	10
X57	CRUM	8	0.853	4	9
X78	CRUM	10	0.499	4	13
X78	CRUM	7	0.336	4	10
X85	CRUM	8	0.945	4	9

Table C 12. Parameter summary for DINA slip parameters across Geweke windows (p1 = p2 = 0.4)

Item	First Mean	Last Mean	Difference	First Variance	Last Variance
X5	0.9050	0.9061	-0.0011	0.0003	0.0003
X6	0.9449	0.9461	-0.0012	0.0001	0.0001
X9	0.9594	0.9603	-0.0009	0.0000	0.0000
X11	0.7166	0.7098	0.0069	0.0012	0.0012
X13	0.8938	0.8921	0.0017	0.0002	0.0001
X15	0.8332	0.8323	0.0008	0.0005	0.0006
X17	0.8098	0.8098	0.0000	0.0007	0.0007
X18	0.8639	0.8631	0.0008	0.0006	0.0006
X19	0.9512	0.9516	-0.0005	0.0001	0.0000
X20	0.8859	0.8882	-0.0022	0.0002	0.0002
X21	0.9669	0.9679	-0.0010	0.0000	0.0000
X22	0.9682	0.9686	-0.0004	0.0000	0.0000
X23	0.9347	0.9355	-0.0009	0.0001	0.0002
X24	0.8164	0.8182	-0.0018	0.0008	0.0009
X25	0.9486	0.9487	-0.0001	0.0000	0.0001
X26	0.9574	0.9568	0.0006	0.0000	0.0000
X27	0.8264	0.8254	0.0010	0.0006	0.0005
X28	0.9533	0.9536	-0.0003	0.0000	0.0001
X29	0.9294	0.9292	0.0002	0.0001	0.0001
X30	0.9655	0.9672	-0.0018	0.0000	0.0000
X31	0.8476	0.8550	-0.0075	0.0008	0.0005
X32	0.2335	0.2333	0.0002	0.0001	0.0001
X33	0.8543	0.8588	-0.0044	0.0002	0.0003
X35	0.9836	0.9804	0.0032	0.0000	0.0000
X39	0.9834	0.9873	-0.0039	0.0000	0.0000
X40	0.9492	0.9460	0.0033	0.0001	0.0001
X41	0.8820	0.8850	-0.0029	0.0002	0.0002
X42	0.9402	0.9381	0.0021	0.0001	0.0002
X43	0.7403	0.7522	-0.0119	0.0014	0.0011
X44	0.9678	0.9672	0.0006	0.0000	0.0000
X45	0.9605	0.9621	-0.0015	0.0000	0.0000
X46	0.9093	0.9116	-0.0023	0.0002	0.0002
X47	0.9240	0.9262	-0.0022	0.0001	0.0001
X48	0.9425	0.9455	-0.0031	0.0000	0.0000
X49	0.9005	0.9031	-0.0026	0.0003	0.0004
X50	0.9903	0.9898	0.0005	0.0000	0.0000
X52	0.9686	0.9689	-0.0003	0.0000	0.0000

Table C12 (continued)

X53	0.9610	0.9643	-0.0033	0.0000	0.0000
X54	0.9172	0.9165	0.0007	0.0002	0.0002
X55	0.9685	0.9644	0.0041	0.0000	0.0001
X56	0.9043	0.9079	-0.0036	0.0001	0.0001
X57	0.9538	0.9546	-0.0008	0.0000	0.0000
X61	0.8400	0.8342	0.0058	0.0003	0.0004
X63	0.9722	0.9707	0.0015	0.0000	0.0000
X64	0.8893	0.8860	0.0033	0.0004	0.0006
X67	0.9065	0.9044	0.0021	0.0001	0.0001
X68	0.8785	0.8825	-0.0040	0.0002	0.0002
X69	0.9217	0.9219	-0.0002	0.0001	0.0001
X70	0.8997	0.9018	-0.0021	0.0001	0.0002
X71	0.9176	0.9189	-0.0013	0.0001	0.0001
X72	0.9043	0.9075	-0.0032	0.0001	0.0001
X73	0.6829	0.6862	-0.0034	0.0008	0.0010
X76	0.8103	0.8165	-0.0062	0.0006	0.0006
X78	0.9167	0.9178	-0.0011	0.0001	0.0001
X79	0.9371	0.9397	-0.0026	0.0001	0.0001
X80	0.7703	0.7663	0.0040	0.0009	0.0010
X81	0.8867	0.8809	0.0058	0.0002	0.0003
X82	0.8523	0.8504	0.0019	0.0003	0.0003
X83	0.9723	0.9691	0.0032	0.0000	0.0001
X84	0.9273	0.9274	0.0000	0.0001	0.0001
X85	0.8058	0.8147	-0.0089	0.0005	0.0006

Table C 13. Parameter summary for DINA guess parameters across Geweke windows (p1 = p2 = 0.4)

Item	First Mean	Last Mean	Difference	First Variance	Last Variance
X5	0.3435	0.3446	-0.0011	0.0001	0.0001
X6	0.6195	0.6210	-0.0015	0.0002	0.0002
X9	0.6780	0.6782	-0.0002	0.0002	0.0002
X11	0.2866	0.2856	0.0010	0.0002	0.0001
X13	0.6145	0.6138	0.0007	0.0002	0.0002
X15	0.3498	0.3503	-0.0005	0.0002	0.0002
X17	0.3535	0.3546	-0.0011	0.0002	0.0002
X18	0.3082	0.3098	-0.0017	0.0002	0.0002
X19	0.5929	0.5903	0.0026	0.0002	0.0002
X20	0.4801	0.4814	-0.0014	0.0002	0.0002
X21	0.5647	0.5744	-0.0097	0.0002	0.0002
X22	0.6121	0.6095	0.0026	0.0002	0.0003
X23	0.4159	0.4150	0.0009	0.0002	0.0002
X24	0.3345	0.3326	0.0019	0.0002	0.0001
X25	0.6290	0.6303	-0.0013	0.0002	0.0002
X26	0.7697	0.7687	0.0010	0.0001	0.0001
X27	0.4007	0.4019	-0.0013	0.0002	0.0002
X28	0.6707	0.6725	-0.0018	0.0002	0.0002
X29	0.4861	0.4857	0.0004	0.0003	0.0002
X30	0.6170	0.6219	-0.0050	0.0002	0.0002
X31	0.3934	0.4004	-0.0070	0.0005	0.0004
X32	0.2335	0.2333	0.0002	0.0001	0.0001
X33	0.5867	0.5911	-0.0045	0.0002	0.0002
X35	0.7644	0.7517	0.0127	0.0004	0.0007
X39	0.6468	0.6719	-0.0251	0.0008	0.0004
X40	0.4773	0.4736	0.0037	0.0003	0.0004
X41	0.5227	0.5258	-0.0031	0.0002	0.0002
X42	0.4121	0.4131	-0.0010	0.0003	0.0004
X43	0.2776	0.2848	-0.0072	0.0002	0.0002
X44	0.6503	0.6497	0.0006	0.0002	0.0002
X45	0.6439	0.6474	-0.0034	0.0002	0.0002
X46	0.3951	0.3952	-0.0001	0.0002	0.0002
X47	0.6524	0.6542	-0.0018	0.0001	0.0001
X48	0.7076	0.7125	-0.0049	0.0001	0.0001
X49	0.3270	0.3298	-0.0028	0.0002	0.0002
X50	0.8921	0.8904	0.0016	0.0001	0.0001
X52	0.6707	0.6720	-0.0013	0.0002	0.0002

Table C13 (continued)

X53	0.7237	0.7315	-0.0078	0.0001	0.0001
X54	0.8085	0.8083	0.0002	0.0002	0.0002
X55	0.8305	0.8221	0.0083	0.0002	0.0002
X56	0.6589	0.6640	-0.0052	0.0002	0.0001
X57	0.7749	0.7761	-0.0012	0.0002	0.0001
X61	0.4447	0.4400	0.0048	0.0001	0.0002
X63	0.8366	0.8336	0.0030	0.0001	0.0001
X64	0.3100	0.3080	0.0020	0.0001	0.0002
X67	0.6247	0.6214	0.0033	0.0001	0.0001
X68	0.5372	0.5410	-0.0038	0.0002	0.0002
X69	0.5457	0.5453	0.0004	0.0002	0.0003
X70	0.5740	0.5772	-0.0031	0.0002	0.0002
X71	0.5311	0.5304	0.0007	0.0001	0.0002
X72	0.5769	0.5808	-0.0039	0.0002	0.0002
X73	0.3429	0.3445	-0.0016	0.0001	0.0002
X76	0.3999	0.4002	-0.0003	0.0002	0.0003
X78	0.6064	0.6073	-0.0010	0.0002	0.0002
X79	0.6080	0.6106	-0.0025	0.0002	0.0002
X80	0.2765	0.2751	0.0014	0.0001	0.0001
X81	0.4948	0.4910	0.0038	0.0002	0.0002
X82	0.4587	0.4572	0.0015	0.0001	0.0001
X83	0.4336	0.4307	0.0030	0.0002	0.0003
X84	0.6909	0.6906	0.0003	0.0001	0.0001
X85	0.4118	0.4173	-0.0055	0.0002	0.0002

Table C 14. Parameter summary for CRUM parameters across Geweke windows (p1 = p2 = 0.4)

Item	Attribute	First Mean	Last Mean	Difference	First Variance	Last Variance
X5	intercept	0.1324	0.1414	-0.0090	0.0051	0.0056
X5	2	1.5023	1.5222	-0.0199	0.0545	0.0471
X5	4	1.2798	1.3383	-0.0584	0.0296	0.0251
X5	7	0.7205	0.6047	0.1158	0.0614	0.0417
X6	intercept	1.0659	1.0640	0.0019	0.0063	0.0060
X6	2	1.1337	1.0321	0.1016	0.0154	0.0208
X6	4	0.8759	0.8925	-0.0166	0.0133	0.0126
X6	6	1.5284	1.7043	-0.1759	0.0293	0.0447
X9	intercept	1.3112	1.2538	0.0574	0.0092	0.0068
X9	2	0.6737	0.5522	0.1215	0.0179	0.0233
X9	4	0.7761	0.7593	0.0168	0.0174	0.0167
X9	6	2.1430	2.2556	-0.1126	0.0454	0.0461
X9	10	0.1428	0.1839	-0.0411	0.0118	0.0149
X11	intercept	-0.7630	-0.7618	-0.0012	0.0041	0.0053
X11	1	0.6870	0.7161	-0.0291	0.0391	0.0404
X11	2	1.7186	1.6314	0.0873	0.0205	0.0246
X13	intercept	0.7081	0.7222	-0.0141	0.0058	0.0071
X13	1	0.9888	0.9167	0.0721	0.0150	0.0138
X13	2	0.6773	0.6296	0.0477	0.0191	0.0162
X13	4	0.2681	0.2984	-0.0303	0.0117	0.0174
X13	8	0.6022	0.7443	-0.1421	0.0395	0.0182
X15	intercept	-0.4141	-0.3749	-0.0392	0.0048	0.0062
X15	1	0.7928	0.8245	-0.0317	0.0467	0.0474
X15	2	2.2496	2.2129	0.0368	0.0278	0.0270
X17	intercept	-0.1884	-0.1744	-0.0140	0.0046	0.0053
X17	1	0.9935	1.0088	-0.0154	0.0463	0.0400
X17	2	0.9496	1.0079	-0.0584	0.0244	0.0255
X17	4	1.1081	1.0557	0.0524	0.0195	0.0202
X18	intercept	-0.2443	-0.2318	-0.0125	0.0056	0.0048
X18	1	0.8096	0.8657	-0.0561	0.0584	0.0772
X18	2	1.6911	1.7178	-0.0267	0.0498	0.0434
X18	4	1.3449	1.2795	0.0654	0.0286	0.0187
X19	intercept	0.9971	1.0051	-0.0080	0.0051	0.0050
X19	1	0.8306	0.7164	0.1142	0.0315	0.0239
X19	5	1.3303	1.2756	0.0547	0.0360	0.0782
X19	6	0.1756	0.1552	0.0204	0.0174	0.0141

Table C14 (continued)

X19	10	1.4558	1.6015	-0.1457	0.0332	0.0794
X20	intercept	0.2541	0.2945	-0.0403	0.0043	0.0033
X20	1	0.2994	0.2401	0.0592	0.0273	0.0232
X20	5	2.6219	2.7089	-0.0871	0.0229	0.0189
X21	intercept	0.8756	0.8594	0.0161	0.0076	0.0076
X21	1	0.5563	0.4957	0.0605	0.0750	0.0541
X21	7	4.6431	4.5317	0.1114	0.0603	0.0656
X22	intercept	0.7821	0.7525	0.0296	0.0063	0.0056
X22	1	0.9405	0.7053	0.2352	0.0415	0.0603
X22	10	4.3423	4.7392	-0.3969	0.0795	0.1254
X23	intercept	0.1453	0.1889	-0.0436	0.0055	0.0049
X23	1	0.2942	0.2567	0.0375	0.0413	0.0353
X23	4	0.4720	0.5410	-0.0690	0.0178	0.0193
X23	5	3.6593	3.5545	0.1048	0.0495	0.0637
X24	intercept	-0.2962	-0.3098	0.0136	0.0057	0.0055
X24	1	0.6257	0.6115	0.0142	0.0587	0.0486
X24	4	1.0622	1.1635	-0.1014	0.0249	0.0201
X24	6	1.3515	1.2235	0.1280	0.0817	0.0704
X24	10	0.4910	0.4356	0.0555	0.0269	0.0276
X25	intercept	0.7242	0.6836	0.0406	0.0073	0.0053
X25	1	0.7727	0.7800	-0.0072	0.0305	0.0267
X25	4	1.5467	1.5722	-0.0255	0.0133	0.0150
X25	6	1.1413	1.0331	0.1082	0.0359	0.0333
X26	intercept	1.9037	1.9017	0.0020	0.0087	0.0070
X26	1	1.1706	1.1539	0.0168	0.0126	0.0112
X26	4	1.0757	1.1356	-0.0599	0.0152	0.0155
X26	5	0.1189	0.1262	-0.0073	0.0067	0.0105
X26	7	0.1799	0.2219	-0.0421	0.0127	0.0153
X26	10	0.5888	0.4906	0.0982	0.0276	0.0351
X27	intercept	-0.0624	-0.0976	0.0352	0.0038	0.0043
X27	1	0.5047	0.4832	0.0215	0.0464	0.0298
X27	4	1.1224	1.1503	-0.0279	0.0095	0.0144
X27	6	0.3886	0.2583	0.1303	0.0501	0.0274
X27	8	0.9990	1.0621	-0.0631	0.0989	0.0420
X28	intercept	1.0159	1.0121	0.0038	0.0059	0.0057
X28	1	0.6419	0.5654	0.0765	0.0217	0.0219
X28	4	1.2189	1.2582	-0.0392	0.0118	0.0119
X28	6	0.2664	0.1535	0.1129	0.0308	0.0145
X28	8	1.1637	1.4139	-0.2502	0.0506	0.0238
X29	intercept	0.4200	0.4468	-0.0268	0.0053	0.0048

Table C14 (continued)

X29	1	1.3584	1.2986	0.0598	0.0500	0.0488
X29	4	1.6397	1.7823	-0.1426	0.0198	0.0163
X29	10	0.9272	0.8858	0.0413	0.0271	0.0279
X30	intercept	0.9038	0.9034	0.0003	0.0057	0.0058
X30	1	0.8513	0.8591	-0.0078	0.0304	0.0299
X30	4	1.4488	1.5586	-0.1098	0.0111	0.0130
X30	6	0.5716	0.3920	0.1796	0.0679	0.0323
X30	8	1.4504	1.4831	-0.0327	0.1689	0.0395
X31	intercept	-1.0275	-1.0595	0.0320	0.0062	0.0058
X31	4	1.8435	1.8187	0.0248	0.0171	0.0181
X32	intercept	-1.1823	-1.1819	-0.0004	0.0020	0.0018
X33	intercept	0.6683	0.6307	0.0376	0.0064	0.0057
X33	4	0.8400	0.7995	0.0405	0.0138	0.0126
X33	6	1.2616	1.2455	0.0161	0.0222	0.0260
X33	7	0.0683	0.0880	-0.0198	0.0034	0.0045
X35	intercept	0.2740	0.2382	0.0359	0.0049	0.0052
X35	4	2.6074	2.5888	0.0186	0.0155	0.0151
X39	intercept	-0.4728	-0.5197	0.0469	0.0042	0.0072
X39	4	2.9122	2.8757	0.0366	0.0246	0.0339
X40	intercept	-0.2460	-0.2753	0.0293	0.0055	0.0041
X40	1	1.9954	1.9817	0.0137	0.0788	0.1419
X40	4	2.3704	2.3654	0.0050	0.0205	0.0200
X41	intercept	0.5413	0.5390	0.0023	0.0045	0.0060
X41	1	1.2164	1.2882	-0.0718	0.0346	0.0262
X41	4	0.9484	0.9975	-0.0491	0.0159	0.0141
X41	5	0.1742	0.1921	-0.0179	0.0105	0.0131
X41	8	1.0930	0.9348	0.1583	0.0569	0.0224
X42	intercept	-0.5318	-0.5596	0.0279	0.0049	0.0047
X42	1	1.3697	1.4914	-0.1217	0.0957	0.0754
X42	4	2.4795	2.3900	0.0895	0.0277	0.0268
X43	intercept	-0.6103	-0.5891	-0.0213	0.0071	0.0074
X43	1	1.6674	1.5331	0.1344	0.1039	0.0815
X43	4	0.9183	0.9212	-0.0028	0.0278	0.0258
X43	7	0.7687	0.8625	-0.0939	0.0225	0.0250
X44	intercept	1.0737	1.1113	-0.0376	0.0072	0.0074
X44	1	2.2508	2.3722	-0.1214	0.0567	0.0947
X44	4	0.9464	0.9762	-0.0297	0.0215	0.0157
X44	5	1.3002	1.3138	-0.0136	0.0203	0.0142
X45	intercept	1.1320	1.1305	0.0015	0.0061	0.0064
X45	1	2.0124	2.0543	-0.0419	0.0391	0.0465

Table C14 (continued)

X45	4	1.3607	1.3550	0.0056	0.0165	0.0175
X45	5	0.8377	0.8527	-0.0150	0.0107	0.0148
X46	intercept	0.0493	0.0378	0.0115	0.0064	0.0047
X46	1	1.4987	1.5625	-0.0638	0.0601	0.0826
X46	4	1.3031	1.3118	-0.0088	0.0194	0.0170
X46	5	1.0486	0.9997	0.0489	0.0250	0.0213
X47	intercept	1.0163	1.0249	-0.0086	0.0060	0.0058
X47	1	0.9679	0.9606	0.0073	0.0145	0.0168
X47	4	0.9291	1.0011	-0.0720	0.0158	0.0169
X47	5	0.8130	0.7603	0.0527	0.0120	0.0092
X48	intercept	1.2128	1.2205	-0.0077	0.0064	0.0075
X48	1	1.0453	1.0112	0.0341	0.0196	0.0158
X48	2	0.3156	0.3854	-0.0698	0.0194	0.0205
X48	4	0.6255	0.6047	0.0208	0.0169	0.0154
X48	8	1.1655	1.1802	-0.0147	0.0251	0.0172
X49	intercept	-0.2478	-0.2328	-0.0150	0.0060	0.0054
X49	1	1.2811	1.3049	-0.0237	0.0714	0.0894
X49	2	1.6881	1.6671	0.0210	0.0372	0.0377
X49	4	1.0742	1.0638	0.0104	0.0276	0.0184
X50	intercept	2.8909	2.8884	0.0025	0.0160	0.0179
X50	1	1.7701	1.8309	-0.0608	0.0149	0.0170
X50	2	0.5915	0.5019	0.0895	0.0258	0.0319
X50	4	0.7109	0.8051	-0.0942	0.0355	0.0164
X50	5	1.0238	0.9712	0.0526	0.0295	0.0301
X52	intercept	1.5971	1.5941	0.0030	0.0088	0.0109
X52	1	1.8574	1.8804	-0.0230	0.0441	0.0357
X52	2	0.5597	0.6472	-0.0875	0.0306	0.0231
X52	4	1.4149	1.4282	-0.0133	0.0226	0.0208
X52	7	0.6964	0.5894	0.1070	0.0188	0.0241
X53	intercept	1.7554	1.7395	0.0158	0.0103	0.0085
X53	1	1.4302	1.4482	-0.0180	0.0171	0.0183
X53	2	0.7579	0.6065	0.1514	0.0207	0.0225
X53	4	0.9937	0.9436	0.0501	0.0220	0.0153
X53	7	0.5970	0.7818	-0.1848	0.0216	0.0197
X54	intercept	1.4670	1.4413	0.0257	0.0052	0.0045
X54	1	2.1505	2.1056	0.0449	0.0142	0.0124
X54	4	0.0731	0.0733	-0.0002	0.0038	0.0034
X55	intercept	1.2017	1.1782	0.0234	0.0061	0.0061
X55	1	2.7572	2.6560	0.1011	0.0249	0.0231
X55	4	0.4754	0.4886	-0.0132	0.0109	0.0118

Table C14 (continued)

X56	intercept	0.6477	0.6541	-0.0064	0.0036	0.0049
X56	1	2.8873	2.8979	-0.0106	0.0300	0.0373
X56	7	0.4382	0.4626	-0.0244	0.0079	0.0075
X57	intercept	1.3871	1.3632	0.0238	0.0070	0.0069
X57	1	2.5588	2.4364	0.1224	0.0236	0.0220
X57	4	0.1099	0.1181	-0.0083	0.0059	0.0083
X57	5	0.2813	0.2220	0.0592	0.0145	0.0138
X57	8	0.2815	0.4600	-0.1786	0.0219	0.0183
X61	intercept	0.0832	0.0995	-0.0163	0.0065	0.0057
X61	2	0.3151	0.3508	-0.0357	0.0261	0.0228
X61	4	0.8844	0.9296	-0.0452	0.0180	0.0195
X61	6	1.4202	1.3263	0.0939	0.0346	0.0351
X63	intercept	2.0271	2.0496	-0.0226	0.0098	0.0079
X63	2	0.2935	0.1761	0.1174	0.0150	0.0117
X63	4	0.5072	0.6447	-0.1374	0.0151	0.0190
X63	5	0.0867	0.0670	0.0197	0.0048	0.0036
X63	6	1.8284	1.8770	-0.0486	0.0179	0.0225
X64	intercept	-0.1727	-0.1621	-0.0106	0.0059	0.0052
X64	2	1.0785	1.1611	-0.0827	0.0342	0.0294
X64	4	1.5221	1.5584	-0.0362	0.0264	0.0234
X64	6	1.5312	1.1761	0.3551	0.2257	0.0762
X67	intercept	0.8811	0.8698	0.0113	0.0050	0.0072
X67	1	0.8059	0.7840	0.0219	0.0189	0.0179
X67	2	0.3836	0.2963	0.0873	0.0188	0.0199
X67	4	0.5127	0.6013	-0.0886	0.0130	0.0200
X67	6	1.0691	1.0787	-0.0096	0.0237	0.0235
X67	10	0.0822	0.0704	0.0118	0.0067	0.0042
X68	intercept	0.3153	0.3552	-0.0399	0.0038	0.0057
X68	1	1.3381	1.3027	0.0354	0.0252	0.0224
X68	2	1.4757	1.5020	-0.0263	0.0128	0.0133
X69	intercept	0.3816	0.3938	-0.0123	0.0053	0.0047
X69	1	1.6339	1.6861	-0.0522	0.0364	0.0420
X69	2	1.8659	1.7915	0.0744	0.0169	0.0137
X70	intercept	0.5045	0.5533	-0.0487	0.0041	0.0053
X70	1	1.2606	1.2386	0.0220	0.0210	0.0214
X70	2	1.6237	1.6657	-0.0420	0.0109	0.0123
X71	intercept	0.5078	0.5192	-0.0113	0.0060	0.0044
X71	1	1.9520	2.0273	-0.0753	0.0519	0.0498
X71	2	1.3658	1.2991	0.0666	0.0232	0.0228
X71	10	0.5097	0.5361	-0.0264	0.0226	0.0212

Table C14 (continued)

X72	intercept	0.6453	0.6325	0.0128	0.0052	0.0061
X72	1	1.5343	1.6075	-0.0732	0.0243	0.0231
X72	2	1.4881	1.4609	0.0272	0.0165	0.0171
X72	4	0.2811	0.2396	0.0415	0.0136	0.0115
X73	intercept	-0.4154	-0.4035	-0.0119	0.0054	0.0043
X73	1	1.1622	1.2542	-0.0920	0.0274	0.0321
X73	2	0.9746	1.0001	-0.0255	0.0333	0.0207
X73	10	0.1867	0.1264	0.0604	0.0166	0.0109
X76	intercept	-0.1496	-0.1850	0.0354	0.0046	0.0055
X76	1	1.0169	1.1052	-0.0883	0.0332	0.0321
X76	4	0.8598	0.8430	0.0169	0.0131	0.0185
X76	10	0.8728	0.8210	0.0519	0.0223	0.0251
X78	intercept	0.8148	0.8027	0.0121	0.0062	0.0041
X78	1	1.4054	1.4524	-0.0470	0.0243	0.0237
X78	4	0.7879	0.7919	-0.0040	0.0153	0.0157
X78	7	0.1824	0.4045	-0.2221	0.0146	0.0362
X78	10	0.7123	0.4547	0.2576	0.0256	0.0278
X79	intercept	0.6237	0.6442	-0.0205	0.0061	0.0043
X79	1	1.7915	1.7949	-0.0033	0.0274	0.0295
X79	5	1.6466	1.6680	-0.0214	0.0150	0.0126
X80	intercept	-0.8067	-0.7955	-0.0113	0.0047	0.0037
X80	1	1.6769	1.7204	-0.0436	0.0963	0.0962
X80	5	1.6524	1.6244	0.0280	0.0232	0.0215
X81	intercept	0.1727	0.2036	-0.0309	0.0042	0.0036
X81	1	1.6434	1.6340	0.0093	0.0380	0.0467
X81	5	1.5850	1.6355	-0.0506	0.0135	0.0135
X82	intercept	0.0855	0.0689	0.0166	0.0040	0.0038
X82	1	1.2360	1.2654	-0.0294	0.0334	0.0281
X82	5	0.7296	0.6966	0.0329	0.0365	0.0314
X82	10	0.8439	0.8240	0.0199	0.0416	0.0376
X83	intercept	-0.7430	-0.7686	0.0256	0.0033	0.0026
X83	1	2.8946	2.8327	0.0619	0.2488	0.2603
X83	6	7.8842	7.3369	0.5473	2.3641	3.3297
X84	intercept	1.0020	1.0179	-0.0159	0.0054	0.0055
X84	1	0.7517	0.7690	-0.0173	0.0149	0.0173
X84	4	0.4231	0.5117	-0.0886	0.0134	0.0143
X84	6	0.8131	0.8849	-0.0718	0.0449	0.0377
X84	7	0.0784	0.0833	-0.0049	0.0044	0.0041
X84	8	0.6244	0.4247	0.1997	0.0385	0.0248
X85	intercept	-0.0696	-0.0638	-0.0058	0.0049	0.0042

Table C14 (continued)

X85	1	0.7585	0.9674	-0.2089	0.0326	0.0343
X85	4	0.7440	0.7593	-0.0153	0.0178	0.0158
X85	8	1.1746	0.7364	0.4382	0.0730	0.0292
X85	10	0.2532	0.3586	-0.1054	0.0160	0.0194

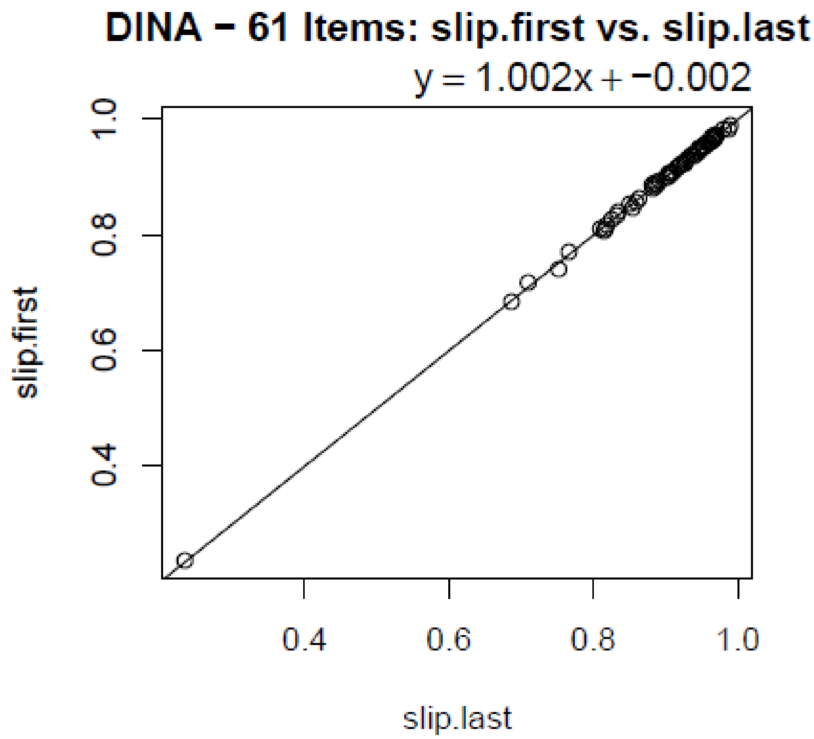

Figure C 1. A regression plot of the reduced DINA slip parameter means from the first and last Geweke windows (p1 = p2 = 0.40)

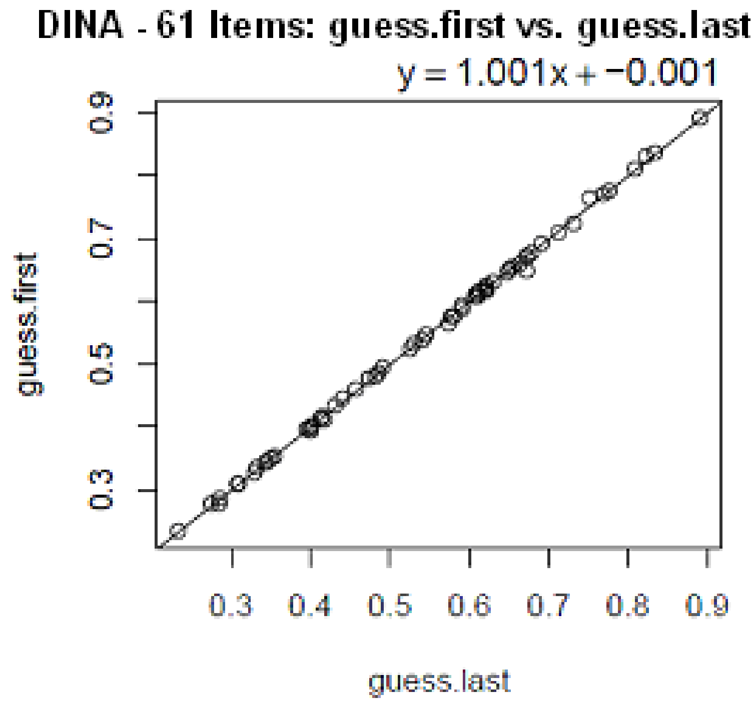

Figure C 2. A regression plot of the reduced DINA guess parameter means from the first and last Geweke windows (p1 = p2 = 0.40)

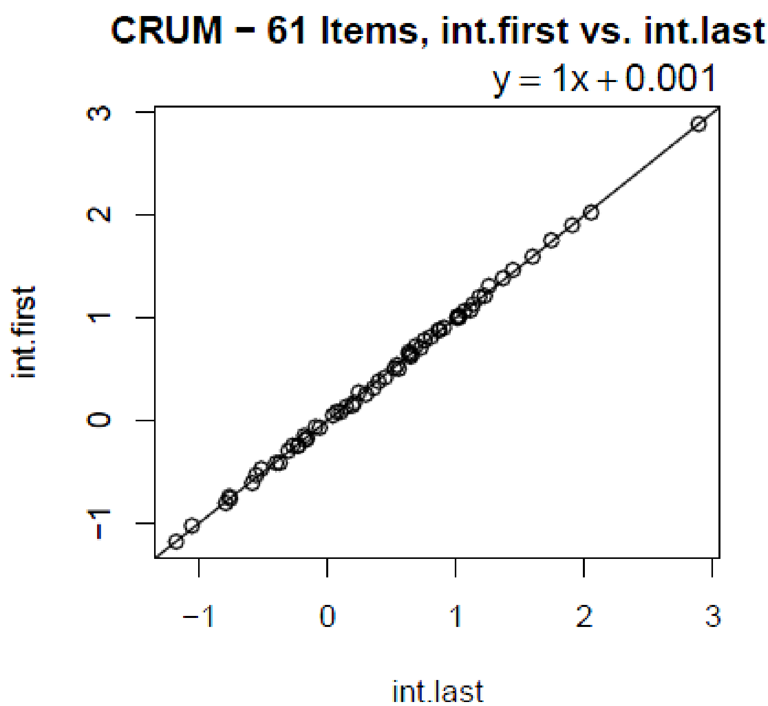

Figure C 3. A regression plot of the reduced CRUM intercept means from the first and last Geweke windows (p1 = p2 = 0.40)

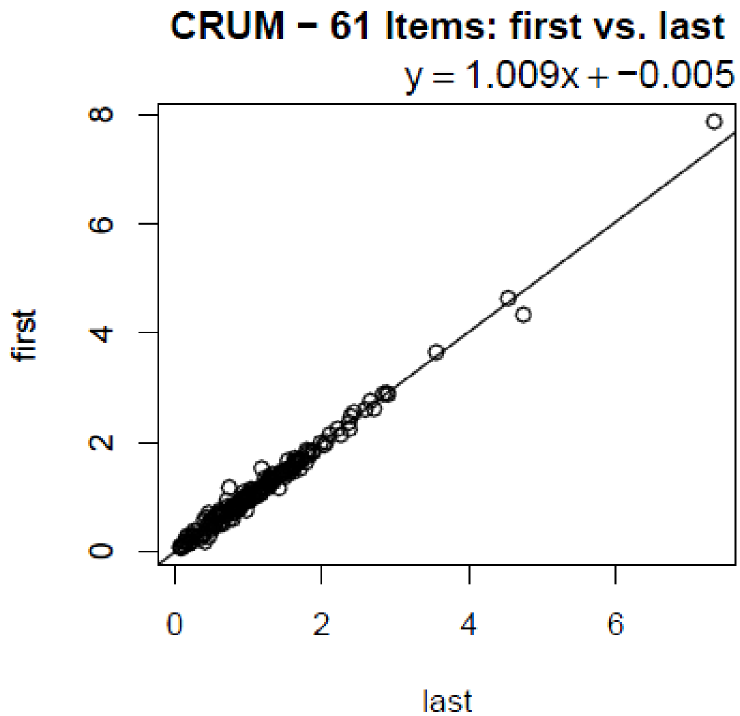

Figure C 4. A regression plot of the reduced CRUM main effects parameter means from the first and last Geweke windows (p1 = p2 = 0.40)

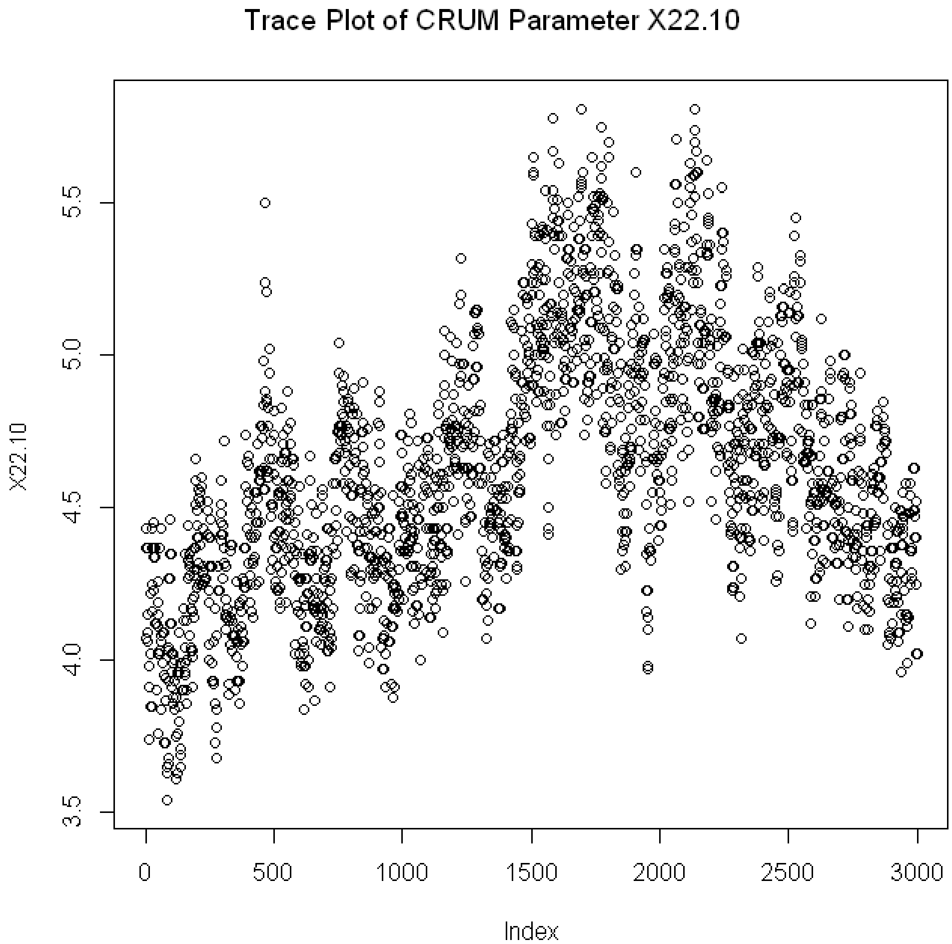

Figure C 5. The trace plot for CRUM main effect for item 2, attribute 10 is unstable, with a marked curvature for steps in the middle of the chain.

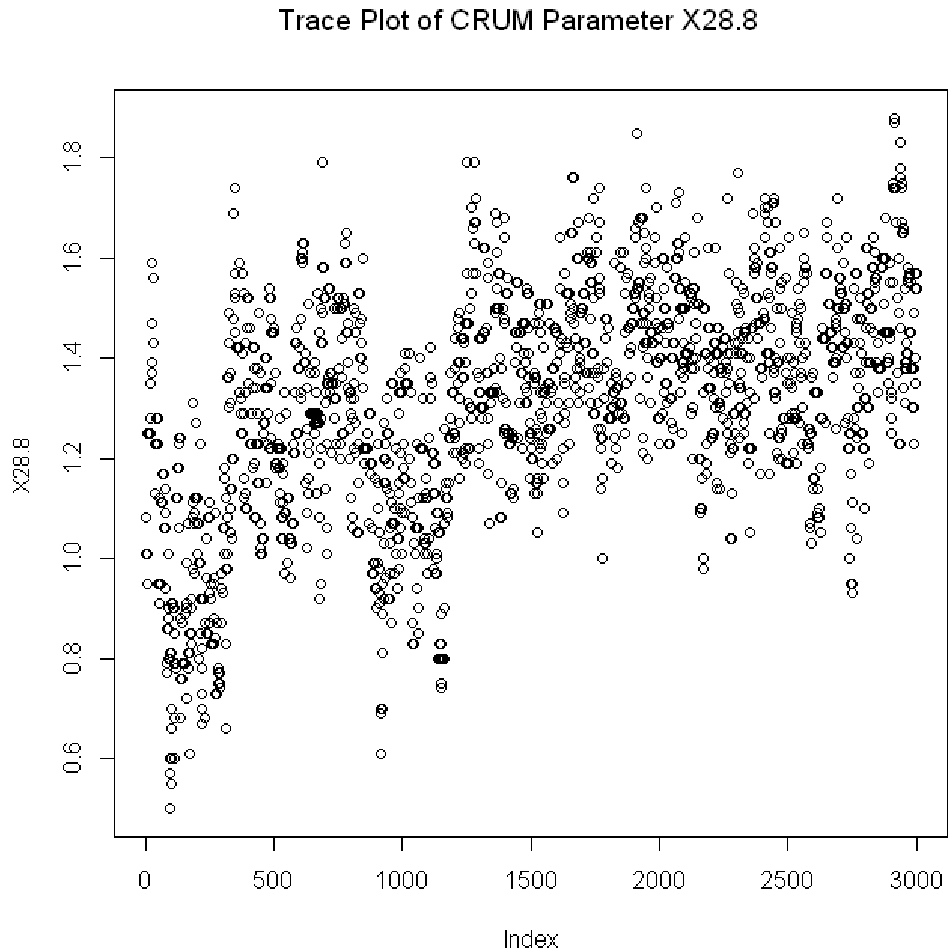

Figure C 6. The trace plot for CRUM main effect for item 28, attribute 28 is unstable, with an increasing trend in the later portion of the chain.

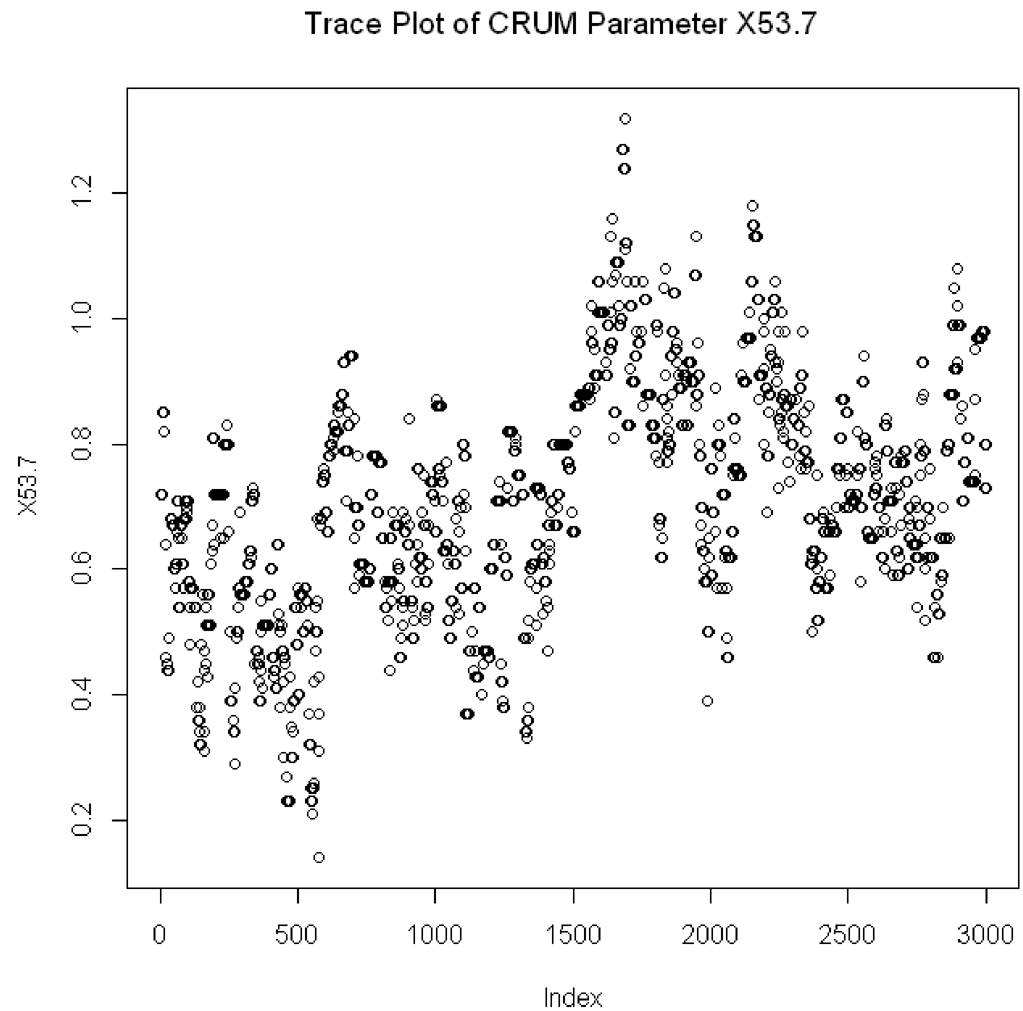

Figure C 7. The trace plot for CRUM main effect for item 53, attribute 7 is unstable, with periodicity and an increasing trend.

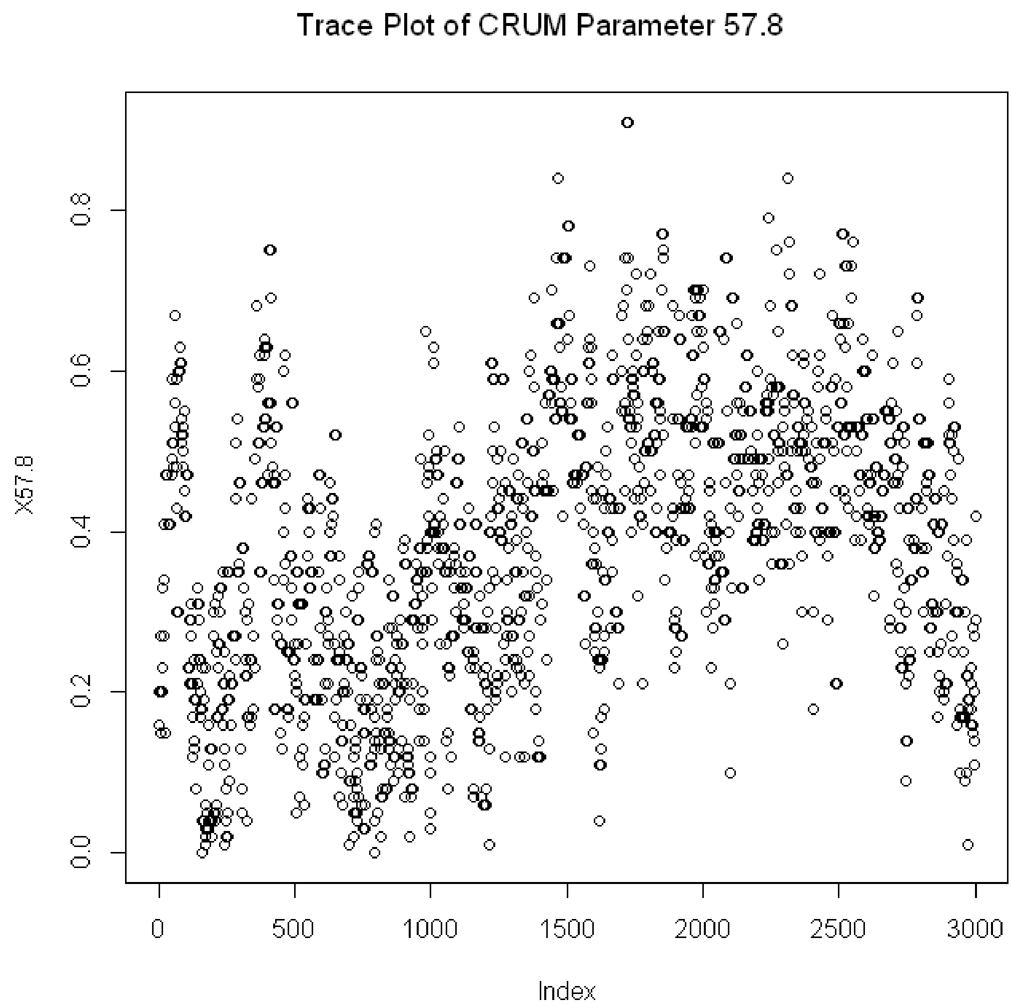

Figure C 8. The trace plot for CRUM main effect for item 57, attribute 8 is unstable, with periodicity over the length of the chain.

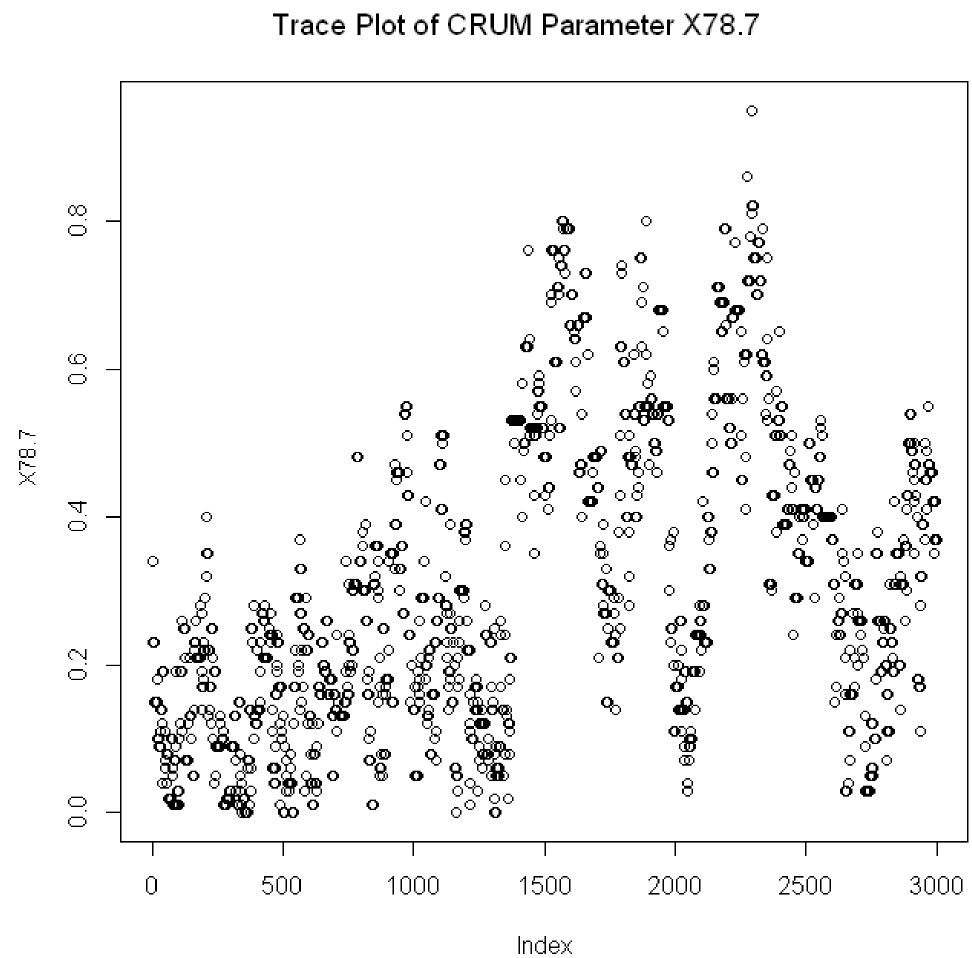

Figure C 9. The trace plot for CRUM main effect for item 78, attribute 7 is unstable, with periodicity and decreasing variance over the length of the chain.

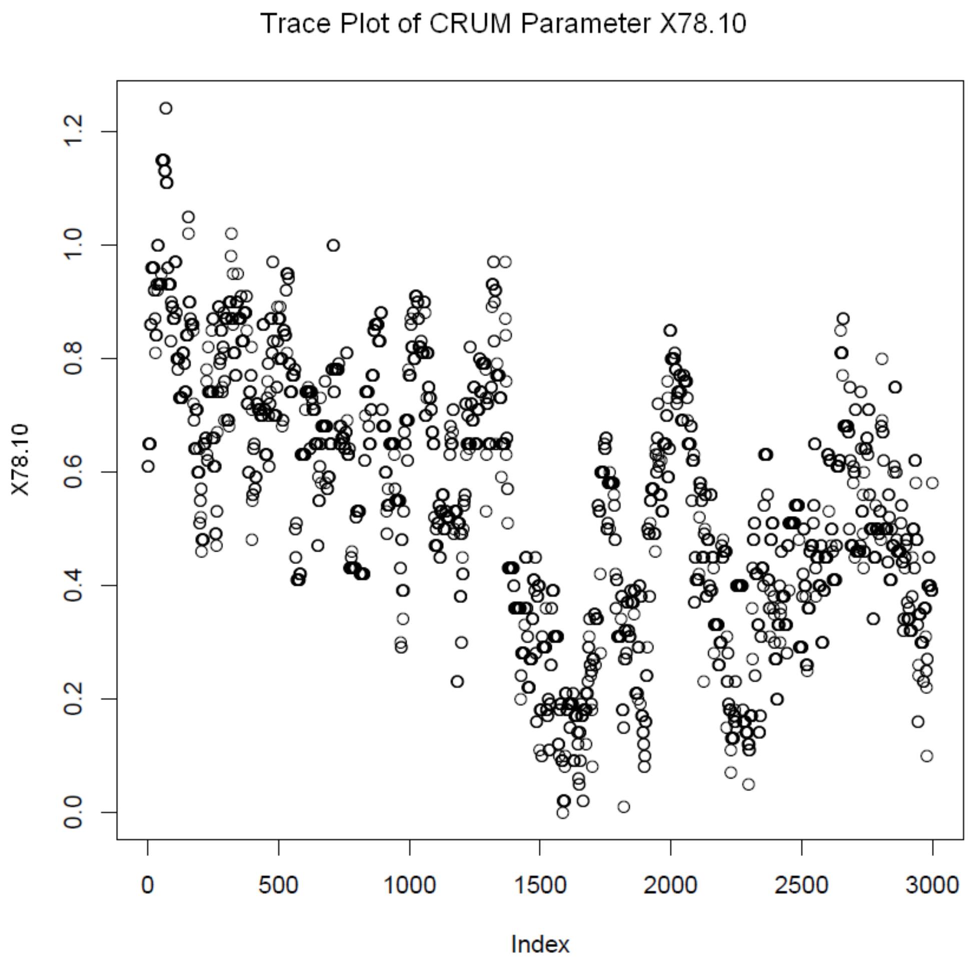

Figure C 10. The trace plot for CRUM main effect for item 78, attribute 10 is unstable, with periodicity and a decreasing trend.

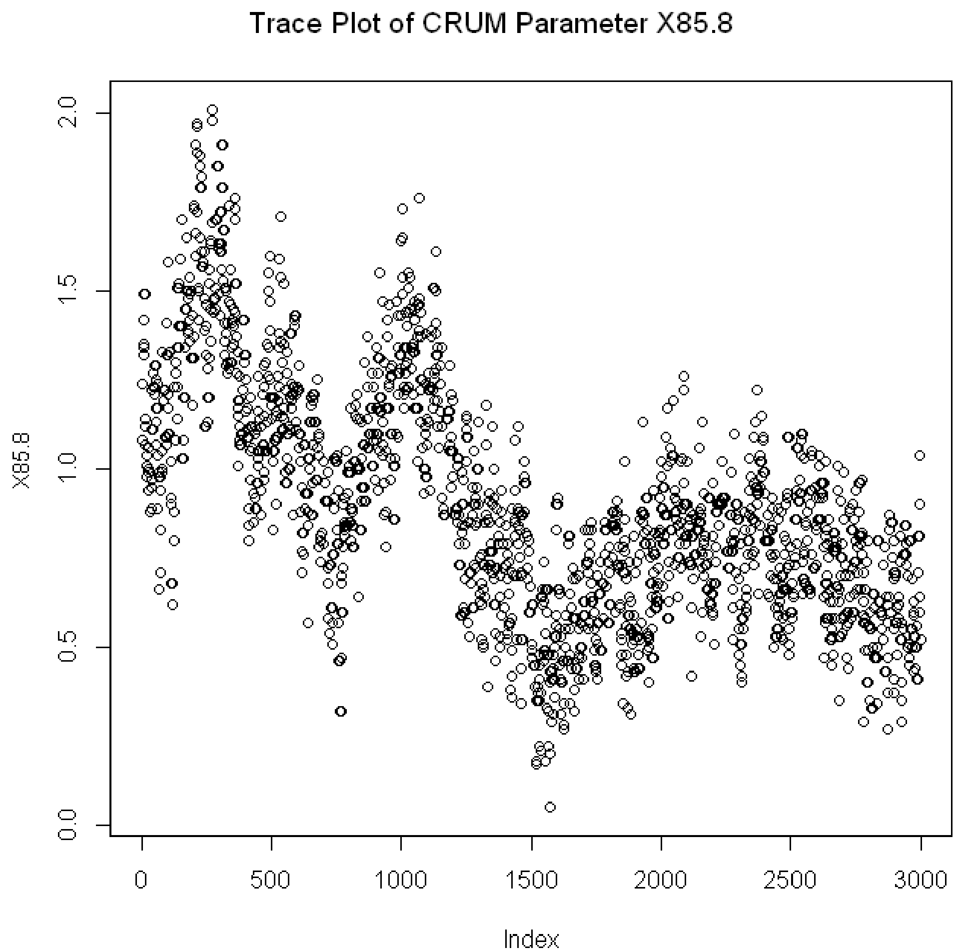

Figure C 11. The trace plot for CRUM main effect for item 85, attribute 8 is unstable, with periodicity and a decreasing trend.

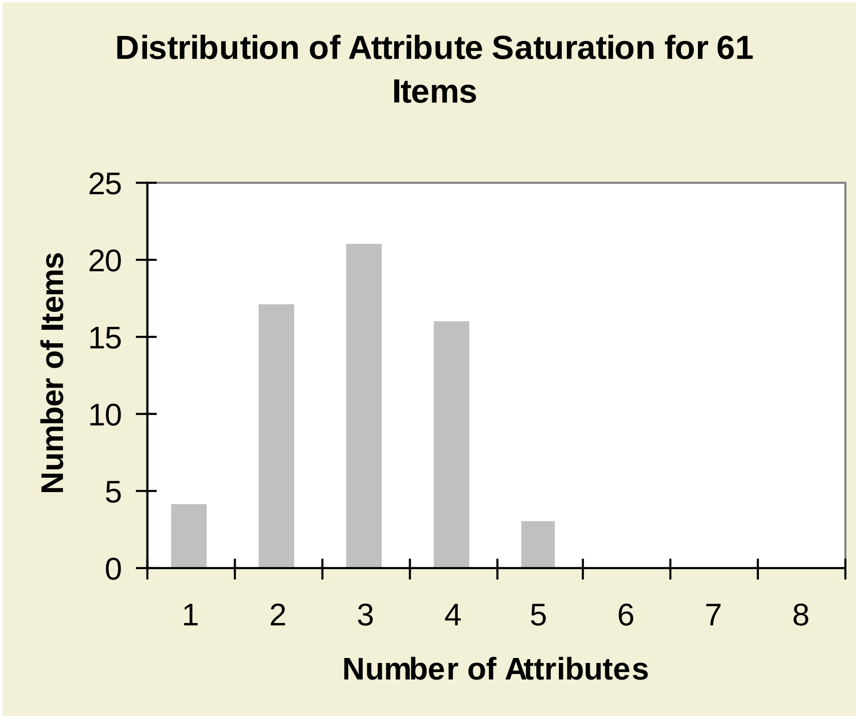

Figure C 12. Distribution of the attribute saturation for the 61 items of the reduced data set.

REFERENCES

Burke, M., & Henson, R. (n.d.) *LCDM user manaul (draft)*. Unpublished user manual.

Embretson, S.E., & Daniel, R.C. (2008). Understanding and quantifying cognitive complexity level in mathematical problem solving items. *Psychological Science Quarterly, 50*(3), 328-344.

Embretson, S. E., & Reise, S. P. (2000). Item response theory for psychologists. Mahwah, NJ: Erlbaum.

Fischer, G. H. (1973). The linear logistic model as an instrument in educational research. *Acta Psychologica, 37*, 359-374.

Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In *Bayesian Statistics 4* (J.M. Bernardo et al., eds), 169-193. Oxford University Press.

Gorin, J.S., & Embretson, S.E. (2006). Item difficulty modeling of paragraph comprehension items. *Applied Psychological Measurement, 30*(5), 394-411.

Haertel, E. (1989). Using restricted latent class models to map the skill structure of achievement items. *Journal of Educational Measurement*, *26*(4), 301-321.

Hartz, S. M. (2002). *A Bayesian framework for the unified model for assessing cognitive abilities: Blending theory with practicality*. (Unpublished doctoral dissertation). University of Illinois at Urbana-Champaign, Urbana, IL.

Henson, R.A., Templin, J.L., & Willse, J.T. (2009). Defining a family of cognitive diagnosis models using log-linear models with latent variables. *Psychometrika, 74*(2), 191-210.

Hornke, L.F. (2002). Item-generation models for higher order cognitive functions. In S. H. Irvine & P. C. Kyllonen (Eds.), *Item generation for test development* (pp. 199-218). Mahwah, New Jersey: Erlbaum.

Huff, K., & Goodman, D.P. (2007). The demand for cognitive diagnostic assessment. In J.P. Leighton & M.J. Geirl (Eds.), *Cognitive diagnostic assessment for education* (19-60). New York, NY: Cambridge University Press.

Junker, B.W., & Sijtsma, K. (2001). Cognitive assessment models with few assumptions, and connections with nonparametric item response theory. *Applied Psychological Measurement, 25*(3), 258-272.

Kansas State Department of Education (2004). *Kansas curricular standards for mathematics education*. Retrieved from http://www.ksde.org/LinkClick.aspx?fileticket=VJee7vha-sQ%3d&tabid=141&mid=8017&forcedownload=true

Lutz, M.E., Embretson, S.E., & Poggio, J. (2010, May). *Assessing the components of mathematics achievement* [Poster]. Presented at the meeting of the American Educational Research Association, Denver, CO.

Mayer, R.E., Larkin, J.H., & Kadane, J.B. (1984). A cognitive analysis of mathematical problem-solving ability. In R. Sternberg (Ed.), *Advances in the psychology of human intelligence.* Hillsdale, NJ: Erlbaum.

Muthén, L.K., & Muthén, B.O. (1998). *Mplus: The comprehensive modeling program for applied researchers.* Los Angeles, CA: Muthén & Muthén.

Rupp, A.A., Templin, J.L., & Henson, R.A. (2010). *Diagnostic measurement: Theory, methods, and applications*. New York, NY: Guilford.

Spiegelhalter, D.J., Best, N.G., Carlin, B.P., and van der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society, 64*(4), 583-639.

Walsh, B. (2004). *Markov Chain Monte Carlo and Gibbs sampling.* Retrieved from http://web.mit.edu/~wingated/www/introductions/mcmc-gibbs-intro.pdf